



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 168293

TO: Jeffrey Parkin
Location: REM-3D39/3C18
Art Unit: 1648
Wednesday, August 03, 2005
Case Serial Number: 10/066506

From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: 571-272-2523

toby.port@uspto.gov

Search Notes

Examiner Parkin,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Toby Port
X22523

This Page Blank (uspto)

STIC-Biotech/ChemLib

160293

From: Parkin, Jeffrey
Sent: Sunday, July 24, 2005 5:36 PM
To: STIC-Biotech/ChemLib
Subject: U.S. Serial No. 10/066,506

Please search **SEQ ID NOS.: 1, 3, 5, and 7** from the aforementioned application (U.S. Serial No. 10/066,506) v. all relevant databases. Place results on both paper and electronic format (i.e., disk, e-mail, etc.).

Thanks!

JSP
AU 1648
REM 3D39
2-0908

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:57:58 ; Search time 9678.51 Seconds
(without alignments)
11274.581 Million cell updates/sec

Title: US-10-066-506A-1

Perfect score: 2252

Sequence: 1 atggcgcttacaggaattt.....tatatggataacggaattc 2252

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2252	100.0	2252	6	AX717713	Sequence
2	1506	66.9	2039	6	AX717723	Sequence
3	1501.4	66.7	11742	14	EBORNA	L11365 Zaire Ebola
4	1499.8	66.6	2298	6	AR404895	Sequence
5	1499.8	66.6	2298	6	AX092102	Sequence
6	1489.4	66.1	2406	14	EVU11033	Zaire Ebola
7	1489.4	66.1	2408	14	EVU23187	Zaire Ebola
8	1489.4	66.1	18959	14	AF086833	Zaire Ebola
9	1489.4	66.1	18959	14	AF272001	Zaire Ebola
10	1489.4	66.1	18959	14	AY142960	Zaire Ebola
11	1487.8	66.1	2408	14	EVU81161	Zaire Ebola
12	1486.2	66.0	18960	14	AF499101	Zaire Ebola
13	1464.6	65.0	18961	14	AY344458	Zaire Ebola
14	1452.6	64.5	2408	14	EVU28077	Zaire Ebola
15	1446.2	64.2	2446	14	EVU77384	Zaire Ebola
16	1444.6	64.1	2173	14	AY058898	Zaire Ebola
17	1435	63.7	2173	14	AY526098	Zaire ebo
18	1435	63.7	2173	14	AY526099	Zaire ebo
19	1435	63.7	2173	14	AY526100	Zaire ebo

20	1435	63.7	2173	14	AY526102	AY526102 Zaire ebo
21	1435	63.7	2173	14	AY526104	AY526104 Zaire ebo
22	1435	63.7	2173	14	AY526105	AY526105 Zaire ebo
23	1433.4	63.7	2173	14	AY526101	AY526101 Zaire ebo
24	753	33.4	2051	6	AX717721	Sequence
25	746.6	33.2	2046	6	AX717719	Sequence
26	742.2	33.0	11460	6	AR279677	Sequence
27	742.2	33.0	19104	14	MVRBPCYC	Z12332 Marburg vir
28	740.6	32.9	19112	14	AY430366	AY430366 Lake Vict
29	740.6	32.9	19112	14	AY430365	AY430365 Lake Vict
30	685.8	30.5	2046	14	AF005735	AF005735 Marburg v
31	685.8	30.5	2948	14	MAVSPAA	X68493 Marburg Vir
32	685.8	30.5	19112	14	MVIRPR	Z29337 Marburg vir
33	675	30.0	19151	14	AY558025	AY558025 Marburg v
34	673	29.9	2046	14	AF005733	AF005733 Marburg v
35	548.2	24.3	2046	6	AX717725	AX717725 Sequence
36	546.6	24.3	2046	6	AX717717	AX717717 Sequence
37	543.4	24.1	2046	14	AF005734	AF005734 Marburg v
38	534.2	23.7	2408	14	EVU28006	U28006 Cote d'Ivoi
39	466.2	20.7	18890	14	AB050936	AB050936 Reston Eb
40	464.6	20.6	2360	14	EVU23152	U23152 Reston Ebol
41	463	20.6	2233	14	AF034645	AF034645 Ebola vir
42	461.4	20.5	18891	14	AF522874	AF522874 Reston Eb
43	459.8	20.4	2033	14	EVU23416	U23416 Reston Ebol
44	456.6	20.3	2033	14	EVU23417	U23417 Reston Ebol
45	428.2	19.0	2261	14	AY344234	AY344234 Sudan ebo

ALIGNMENTS

RESULT 1
AX717713
LOCUS AX717713 2252 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 1 from Patent WO02079239.
ACCESSION AX717713
VERSION AX717713.1 GI:29890725
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 1 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source
1..2252
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="chimeric molecule between Ebola virus Zaire Mayinga strain Glycoprotein 1 and Marburg virus strain Musoke Glycoprotein 2"

ORIGIN	Query Match	100.0%;	Score 2252;	DB 6;	Length 2252;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2252;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGGGCGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAGAGGACATCATCTTTT	60		
Db	1	ATGGGCGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAGAGGACATCATCTTTT	60		
Qy	61	CTTTGGGTAATATCTCTTTTCCAAAGAACATTTTCCATCCCATCTTGAGTCATCCCAAT	120		
Db	61	CTTTGGGTAATATCTCTTTTCCAAAGAACATTTTCCATCCCATCTTGAGTCATCCCAAT	120		
Qy	121	AGCATATTACAGGTTAGTGTGCGACAACTAGTTTGTGTCGACAACTGTCATCCACA	180		
Db	121	AGCATATTACAGGTTAGTGTGCGACAACTAGTTTGTGTCGACAACTGTCATCCACA	180		
Qy	181	AATCAATTGATGATGATGGGAACTGTCGAAGGGAAATGAGTGGCAACTGACGTGCA	240		

181 ATCAATTCAGATCAGTTGGAATCTCGAAGGAATGAGTGGCAACTGACGTGCA 240
241 TCTGCAACTAAAGATGGGCTTTCAGTCCGGTGTCCACCAAGGTGCTCAATATGAA 300
241 TCTGCACATAAAGATGGGCTTTCAGTCCGGTGTCCACCAAGGTGCTCAATATGAA 300
301 GCTGTGAATGGCTGAAACTGCTACAATCTTGAATCAAAACCACTGACGGGAGTGAG 360
301 GCTGTGAATGGCTGAAACTGCTACAATCTTGAATCAAAACCACTGACGGGAGTGAG 360
361 TGTCTACAGACGCGCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGACAAA 420
361 TGTCTACAGACGCGCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGACAAA 420
421 GTATCAGGAACGGGACCGGTGCGGAGACTTTCCTCCATAAAGAGGCTCTTCTTC 480
421 GTATCAGGAACGGGACCGGTGCGGAGACTTTCCTCCATAAAGAGGCTCTTCTTC 480
481 CTGTATGATCGACTTGTCTCCAGTTATCTACCGAGGAACGACTTTCGCTGAAGTGTG 540
481 CTGTATGATCGACTTGTCTCCAGTTATCTACCGAGGAACGACTTTCGCTGAAGTGTG 540
541 GTTGCATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTCAGCTCACACCCCTTGAGA 600
541 GTTGCATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTCAGCTCACACCCCTTGAGA 600
601 GAGCGGTCAATGCAACGGAGACCGGTCTAGTGGCTACTATCTACCAATATAGATAT 660
601 GAGCGGTCAATGCAACGGAGACCGGTCTAGTGGCTACTATCTACCAATATAGATAT 660
661 CAGGCTACCGGTTTGGAAACCAATGAGA CAGAGTACTTCTCGAGGTTGACAATTTGACC 720
661 CAGGCTACCGGTTTGGAAACCAATGAGA CAGAGTACTTCTCGAGGTTGACAATTTGACC 720
721 TAGCTCAACTTGAATCAAGATTCA CACA CAGTTTCTGCTCAGCTGAATGAGACAATA 780
721 TAGCTCAACTTGAATCAAGATTCA CACA CAGTTTCTGCTCAGCTGAATGAGACAATA 780
781 TATCAAGTGGGAAGAGGACCAATCCACGGGAACCAATTTTGAAGGTCAACCCGAA 840
781 TATCAAGTGGGAAGAGGACCAATCCACGGGAACCAATTTTGAAGGTCAACCCGAA 840
841 ATTGATACAACATCGGGAGTGGGCTTCTGGGAACTTAAAGAACTCTACTAGAAA 900
841 ATTGATACAACATCGGGAGTGGGCTTCTGGGAACTTAAAGAACTCTACTAGAAA 900
901 ATTGAGTGAAGAGTTGTTCTTTCAGTTGATCAAA CCGGACCAAAACATCATGTGTT 960
901 ATTGAGTGAAGAGTTGTTCTTTCAGTTGATCAAA CCGGACCAAAACATCATGTGTT 960
961 CAGAGTCCGGCGGAATCTTCCGACCCAGGACCAACACAACTGAAGACCAAAA 1020
961 CAGAGTCCGGCGGAATCTTCCGACCCAGGACCAACACAACTGAAGACCAAAA 1020
1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1080
1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1080
1081 GCAGTGTGCAATCAACACCTTGCACAAATCTCCA CAGGTCGCCCAATCCCTCAACCC 1140
1081 GCAGTGTGCAATCAACACCTTGCACAAATCTCCA CAGGTCGCCCAATCCCTCAACCC 1140
1141 AAACCAAGTCCGGACCAACGACCCCAATATAC CCGGTGTAATACTTGACATCTCTGAG 1200
1141 AAACCAAGTCCGGACCAACGACCCCAATATAC CCGGTGTAATACTTGACATCTCTGAG 1200
1201 GCAACTCAAGTTGAACAACTCAACCGAGAA CAGACACGACAGCACGCTCCGACACT 1260
1201 GCAACTCAAGTTGAACAACTCAACCGAGAA CAGACACGACAGCACGCTCCGACACT 1260
1261 CCTCTTGCCA CAGACCGGACCGGAC CCCCCAAAGAGAGAA CACCAACGACGAGGAGGC 1320

1261 CCCTCTGCCACGACGCGACCGGACCCCCAAAGCAGAGAA CACAACACGACGAGAGGC 1320
1321 ACTGACTTCTGGACCCCGCCACCAACAAGTCCCAAAAC CACAGCGAGACCCGCTGGC 1380
1321 ACTGACTTCTGGACCCCGCCACCAACAAGTCCCAAAAC CACAGCGAGACCCGCTGGC 1380
1381 AACAAACAATCTCATCAACCAAGATACCGGAGAAGAGT GCGCAGCAGCGGGAAGCTAGGC 1440
1381 AACAAACAATCTCATCAACCAAGATACCGGAGAAGAGT GCGCAGCAGCGGGAAGCTAGGC 1440
1441 TTATTAACAATACTATTTGCTGGAGTCCAGGACTGATCA CAGGCGGGAAGAACTCGA 1500
1441 TTATTAACAATACTATTTGCTGGAGTCCAGGACTGATCA CAGGCGGGAAGAACTCGA 1500
1501 CGATCGATCTCTGGAGGGAAGGCGACATGTTCCCTTTT CTGGATGGGTTAATAATGCT 1560
1501 CGATCGATCTCTGGAGGGAAGGCGACATGTTCCCTTTT CTGGATGGGTTAATAATGCT 1560
1561 CCAATTGATTTTGACCCAGTTCCAAATACAAAACAATCT TGAATAATCTCTAGTTCT 1620
1561 CCAATTGATTTTGACCCAGTTCCAAATACAAAACAATCT TGAATAATCTCTAGTTCT 1620
1621 GGTGCTCGGCTGAGGAAGATCAACATGCTCCCTCCCAAT TTAGTTAACTTTACTTTAT 1680
1621 GGTGCTCGGCTGAGGAAGATCAACATGCTCCCTCCCAAT TTAGTTAACTTTACTTTAT 1680
1681 TTTCTTAATATAATAGAGAACATGCTCTACTCTGAGAAA ATGAGAATGATTTGATGCA 1740
1681 TTTCTTAATATAATAGAGAACATGCTCTACTCTGAGAAA ATGAGAATGATTTGATGCA 1740
1741 GAGTTAAGAAATTTGGAGCGTT CAGGAGGATGACCTGG CCGCAGGGCTCAGTTGGATACCG 1800
1741 GAGTTAAGAAATTTGGAGCGTT CAGGAGGATGACCTGG CCGCAGGGCTCAGTTGGATACCG 1800
1801 TTTTGTGCGCTGGAAATTTGAAGGACTTTACATGCTGT GTTTAAATTTAAAAATCAAAACAT 1860
1801 TTTTGTGCGCTGGAAATTTGAAGGACTTTACATGCTGT GTTTAAATTTAAAAATCAAAACAT 1860
1861 TTGGTCTCAGGTTGAGGCGCTAGCCAATCAAACTGCAAA ATCCTTGGAACTCTTATTG 1920
1861 TTGGTCTCAGGTTGAGGCGCTAGCCAATCAAACTGCAAA ATCCTTGGAACTCTTATTG 1920
1921 AGAGTCACAACCTGAGGAAGAACATTTCTCTTAATCAAT AGATAGATGCTATTGACTTTCTA 1980
1921 AGAGTCACAACCTGAGGAAGAACATTTCTCTTAATCAAT AGATAGATGCTATTGACTTTCTA 1980
1981 CTCACAAGATGGGAGGAACATGCAAGTCTTGGACCTGAT TGTGCACTCGGATAGAA 2040
1981 CTCACAAGATGGGAGGAACATGCAAGTCTTGGACCTGAT TGTGCACTCGGATAGAA 2040
2041 GACTTGTCCAAAATATTTTCAGAGCAAAATTGACCAAAAT TAAAAAGGACGAAACAAAAGAG 2100
2041 GACTTGTCCAAAATATTTTCAGAGCAAAATTGACCAAAAT TAAAAAGGACGAAACAAAAGAG 2100
2101 GGGACTGTGTGGGTCTCGGTGGTAAATGGTGGACA TCCGACTGGGGTCTTCTTACTAAC 2160
2101 GGGACTGTGTGGGTCTCGGTGGTAAATGGTGGACA TCCGACTGGGGTCTTCTTACTAAC 2160
2161 TTGGGCATTTTGTCTACTATATATCCATAGCTGTCTTGA TTGCTCTATCCGTATTTGCTGT 2220
2161 TTGGGCATTTTGTCTACTATATATCCATAGCTGTCTTGA TTGCTCTATCCGTATTTGCTGT 2220
2221 ATCTTTTACTAAATATATCGGATAACGGAATTC 2252
2221 ATCTTTTACTAAATATATCGGATAACGGAATTC 2252

RESULT 2
AX717723
LOCUS
DEFINITION
ACCESSION
VERSION

AX717723 2039 bp DNA linear PAT 15-APR-2003
Sequence 11 from Patent WO02079239.
AX717723
AX717723.1 GI:29890730

KEYWORDS Zaire ebolavirus (ZEBOV)
SOURCE Zaire ebolavirus
ORGANISM Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 11 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES Location/Qualifiers
1..2039
/organism="Zaire ebolavirus"
/mol_type="unassigned DNA"
/db_xref="taxon:186538"
/note="chimeric molecule between Ebola virus Glycoprotein
1 and Ebola virus Glycoprotein 2"
ORIGIN
Query Match 66.9%; Score 1506; DB 6; Length 2039;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGGGCGTTACAGGAATTCGAGTTACCTCGTGATCGATTCAAGAGGACATCTTTT 60
Db 1 ATGGGCGTTACAGGAATTCGAGTTACCTCGTGATCGATTCAAGAGGACATCTTTT 60
Qy 61 CTTTGGGTAATATCTTTTCCAAAGAAACATTTTCCATCCCACTTGGAGTCATCCCAAT 120
Db 61 CTTTGGGTAATATCTTTTCCAAAGAAACATTTTCCATCCCACTTGGAGTCATCCCAAT 120
Qy 121 AGCATTACAGGTTAGTGATGTGACAACTAGTTTGTGCTGACAAACTGTGATCCACA 180
Db 121 AGCATTACAGGTTAGTGATGTGACAAACTAGTTTGTGCTGACAAACTGTGATCCACA 180
Qy 181 AATCAATTGAGATCAGTTGGATCTGAAAGGAAATGAGTGGCAACTGAGTGCCA 240
Db 181 AATCAATTGAGATCAGTTGGATCTGAAAGGAAATGAGTGGCAACTGAGTGCCA 240
Qy 241 TCTGCAACTAAAGATGGGCTTCAGGTCGCGTGTCCACCAAAAGGTGGTCAATATGAA 300
Db 241 TCTGCAACTAAAGATGGGCTTCAGGTCGCGTGTCCACCAAAAGGTGGTCAATATGAA 300
Qy 301 GCTGTGTAATGGCTGAAAACTGCTACAACTTTGAAATCAAAAACTGACGGAGTGAG 360
Db 301 GCTGTGTAATGGCTGAAAACTGCTACAACTTTGAAATCAAAAACTGACGGAGTGAG 360
Qy 361 TGTCTACAGAGCGCCAGAGCGGATTCGGGGCTTCCCGGTCGCGGTATGTGCACAAA 420
Db 361 TGTCTACAGAGCGCCAGAGCGGATTCGGGGCTTCCCGGTCGCGGTATGTGCACAAA 420
Qy 421 GTATCAGGAACGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 421 GTATCAGGAACGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Qy 481 CTGTATGATCGACTTGTTCACAGTTATCTACCGAGAAACGACTTTCGCTGAAGGTGTC 540
Db 481 CTGTATGATCGACTTGTTCACAGTTATCTACCGAGAAACGACTTTCGCTGAAGGTGTC 540
Qy 541 GTTGCAATTCGATCTGCTCCCAAGCTAAGAGGACTTCTCAGCTACACCCCTTGAGA 600
Db 541 GTTGCAATTCGATCTGCTCCCAAGCTAAGAGGACTTCTCAGCTACACCCCTTGAGA 600
Qy 601 GAGCCGGTCAATGCAACCGAGGACCGCTCTAGTGGCTACTATTCTACCAATATAGATAT 660
Db 601 GAGCCGGTCAATGCAACCGAGGACCGCTCTAGTGGCTACTATTCTACCAATATAGATAT 660
Qy 661 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTTGTCGAGGTTGACAAATTGACC 720
Db 661 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTTGTCGAGGTTGACAAATTGACC 720
Qy 721 TAGCTCCAACTTGAATCAAGATTTCACCAACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 721 TAGCTCCAACTTGAATCAAGATTTCACCAACAGTTTCTGCTCCAGCTGAATGAGACAATA 780

Db 721 TAGCTCCAACTTGAATCAAGATTTCACCAACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Qy 781 TATACAAAGTGGGAAAAGAGCAATATACACGGGAAAACCTAATTTTGAAGGTCAACCCGAA 840
Db 781 TATACAAAGTGGGAAAAGAGCAATATACACGGGAAAACCTAATTTTGAAGGTCAACCCGAA 840
Qy 841 ATTGATACAAATCGGGAGTGGCTTCTGGGAACTAATAAAACCTCAGTACAGAAA 900
Db 841 ATTGATACAAATCGGGAGTGGCTTCTGGGAACTAATAAAACCTCAGTACAGAAA 900
Qy 901 ATTGCGAGTCAAGAGTGTCTTTTTCACAGTTGTATCAAAACGAGGCAAAACATCAGTGT 960
Db 901 ATTGCGAGTCAAGAGTGTCTTTTTCACAGTTGTATCAAAACGAGGCAAAACATCAGTGT 960
Qy 961 CAGAGTCCGGCGGAACTTTTTCGACCCAGGGAACCAACAACTGAGAGACACAAA 1020
Db 961 CAGAGTCCGGCGGAACTTTTTCGACCCAGGGAACCAACAACTGAGAGACACAAA 1020
Qy 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAGTGCACAGTCAAGGAAGGAAAGCT 1080
Db 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAGTGCACAGTCAAGGAAGGAAAGCT 1080
Qy 1081 GCAGTGTGCAATCTTAACAACCTTTGCAACATCTCCACGAGTCCCAATCCTCACAAC 1140
Db 1081 GCAGTGTGCAATCTTAACAACCTTTGCAACATCTCCACGAGTCCCAATCCTCACAAC 1140
Qy 1141 AAACAGGTCGGGACACAGACCCATTAATACACCCGTGTATAAATTTGACATCTCTGAG 1200
Db 1141 AAACAGGTCGGGACACAGACCCATTAATACACCCGTGTATAAATTTGACATCTCTGAG 1200
Qy 1201 GCAACTCAAGTTGAAACCAATCAACCGAGAACAGCAACGACAGCAGCCTCCGACACT 1260
Db 1201 GCAACTCAAGTTGAAACCAATCAACCGAGAACAGCAACGACAGCAGCCTCCGACACT 1260
Qy 1261 CCCTCTGCCAGACCGGACCGGACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC 1320
Db 1261 CCCTCTGCCAGACCGGACCGGACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC 1320
Qy 1321 ACTGACTTCTGGACCCGCCACCAACAAAGTCCCAAAACCAACGAGAGACCGCTGGC 1380
Db 1321 ACTGACTTCTGGACCCGCCACCAACAAAGTCCCAAAACCAACGAGAGACCGCTGGC 1380
Qy 1381 AACCAACACTCATCAAGATACCGGAGAGAGAGTCCAGCAGCGGAGAGCTAGGC 1440
Db 1381 AACCAACACTCATCAAGATACCGGAGAGAGAGTCCAGCAGCGGAGAGCTAGGC 1440
Qy 1441 TTAATTACCAATATCTTTTGTGGAGTCCGAGGACTGATCAAGCGGAGAGAACTCGA 1500
Db 1441 TTAATTACCAATATCTTTTGTGGAGTCCGAGGACTGATCAAGCGGAGAGAACTCGA 1500
Qy 1501 CGATCG 1506
Db 1501 CGATCG 1506
RESULT 3
EBORNA 11742 bp ss-RNA linear VRL 28-AUG-2002
LOCUS Zaire Ebola virus nucleoprotein, polymerase complex protein (vp35),
DEFINITION matrix protein (vp40), glycoprotein (gp), minor nucleoprotein
(vp30), and membrane-associated structural protein (vp24)s,
complete cds; and polymerase (POLYMERASE), partial cds.
ACCESSION L11365
VERSION L11365
KEYWORDS L11365.1 GI:2522270
ORGANISM Zaire ebolavirus (ZEBOV)
SOURCE Viruses; ssRNA negative-strand viruses; Mononegavirales;
REFERENCE 1 (bases 1 to 11742)
AUTHORS Sanchez, A., Kiley, M.P., Holloway, B.P., McCormick, J.B. and
Auperin, D.D.
TITLE The nucleoprotein gene of Ebola virus: cloning, sequencing, and in

misc_feature	7792..7869 /gene="gp" /note="encodes immunosuppressive motif"
misc_feature	7987..8055 /gene="gp" /note="encodes transmembrane domain"
gene	8291..9741 /gene="vp30"
mRNA	8291..9741 /gene="vp30" /product="minor nucleoprotein" /note="putative" /citation=[2] 8511..9293 /gene="vp30" /note="putative" /citation=[2] /codon_start=1 /product="minor nucleoprotein" /protein_id="AAB81005.1" /db_xref="GI:290393" /translation="MEASERGRPAAHQSRDCHDHVVARSSSRENYRGEYQRQSR ASORVPVTGPHKRVBLTVPPAPKDCPTLKGFLCDSFCKKHQLSLTDRELL LIARKTGSVEQNLNITAPDKSRLANPTADDFOEBEGPKITLTLLKTAHWAQDIR TIEDSKUALLTCVAMTRFKSPKSQSLSLCETHLRREGLGQDAEPVLVEYQRLHSDK GSGFEALWQMDPRQSLIMFITPALNALQLPCESSAALFQG" /product="intergenic region" /citation=[2] /function="separates vp30 gene and vp24 gene" /evidence=experimental 9884..11518 /gene="vp24" 9884..11518 /gene="vp24" /product="membrane-associated structural protein" /citation=[2] /evidence=experimental
gene	Query Match 66.7%; Score 1501.4; DB 14; Length 11742; Best Local Similarity 99.9%; Pred. No. 0; Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
mRNA	Qy 1 ATGGCGCTTACAGGAATATTGCAGTTACTCGTGATCGATTCGAAGAGGACATCATCTCTTT 60 Db 6041 ATGGCGGTACAGGAATATTGCAGTTACTCGTGATCGATTCGAAGAGGACATCATCTCTTT 6100 Qy 61 CTTTGGGTAAATTATCCCTTTTCCAAGAACAATTTTCCATCCCACTTGAGAGTCATCCCAAT 120 Db 6101 CTTTGGGTAAATTATCCCTTTTCCAAGAACAATTTTCCATCCCACTTGAGAGTCATCCCAAT 6160 Qy 121 AGCACATTACAGGTTAGTAGTGTGCAACAATACTAGTTTGTGCTGACAACCTGTCATCCACA 180 Db 6161 AGCACATTACAGGTTAGTAGTGTGCAACAATACTAGTTTGTGCTGACAACCTGTCATCCACA 6220 Qy 181 AATCAATTGAGATCAGTTGGACTGAACTCTCGAAGGGGAATGGAGTGGCAA CTGACGTGCCA 240 Db 6221 AATCAATTGAGATCAGTTGGACTGAACTCTCGAAGGGGAATGGAGTGGCAA CTGACGTGCCA 6280 Qy 241 TCTGCAATATAAGATGGGCTTCAGTCCGGTGTCCCACAAAAGTGTGTCATTTATGAA 300 Db 6281 TCTGCAATATAAGATGGGCTTCAGTCCGGTGTCCCACAAAAGTGTGTCATTTATGAA 6340 Qy 301 GCTGCTGAATGGCGCTGAATACTGTAACATCTGAATCAAAAACCTGACGGGAGTGAG 360 Db 6341 GCTGCTGAATGGCGCTGAATACTGTAACATCTGAATCAAAAACCTGACGGGAGTGAG 6400 Qy 361 TGCTCTACAGACGGCCAGACGGGATTCGGGGCTTCCCCGGGTCCCGGTATGTGCACAAA 420 Db 6401 TGCTCTACAGACGGCCAGACGGGATTCGGGGCTTCCCCGGGTCCCGGTATGTGCACAAA 6460 Qy 421 GTATCAGGAACGGGACCCTGTGTCGGAGACTTTGCCCTTCACATAAAGAGGGTGCTTTCTTC 480

6461	GTATCAGGAACGGGACCGTGTGCCGAGACTTTGSCCTTCCATAAAGAGGGTGCTTTCTTC	6522		
QY	CTGTATGATCGACTTGTCTTCCACAGTTTATCTACCGAGGAACGACTTTTCGTCTGAAGGTGTC	540		
	6521	CTGTATGATCGACTTGTCTTCCACAGTTTATCTACCGAGGAACGACTTTTCGTCTGAAGGTGTC	6580	
QY	541	GTTGCAATTTCTGATATCTGCCCCAAGCTTAAGAAGCACTTTCTTACGCTCACAACCCCTTGAGA	600	
Db	6581	GTTGCAATTTCTGATATCTGCCCCAAGCTTAAGAAGCACTTTCTTACGCTCACAACCCCTTGAGA	6640	
	QY	601	GAGCCGGTCAATCGAACGGAGGACCGCTCTAGTGGCTACTATTCTACCACAAATAGATAT	660
	Db	6641	GAGCCGGTCAATCGAACGGAGGACCGCTCTAGTGGCTACTATTCTACCACAAATAGATAT	6700
QY		661	CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTTGTCGAGGTTGACAAATTTTGACC	720
Db		6701	CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTTGTCGAGGTTGACAAATTTTGACC	6760
	QY	721	TAGCTCAACTTGAATCAAGATTCAACCAACAGTTTCTGCTCCAGCTGAATAGACAATA	780
	Db	6761	TAGCTCAACTTGAATCAAGATTCAACCAACAGTTTCTGCTCCAGCTGAATAGACAATA	6820
QY		781	TATACAGTGGGAAAAGGAGCAATACACGGGAAAACATAATTTTGGAGGTCACACCCGAA	840
Db		6821	TATACAGTGGGAAAAGGAGCAATACACGGGAAAACATAATTTTGGAGGTCACACCCGAA	6880
	QY	841	ATTGATACAAATCGGGAGTGGGCTTCTGCGAAACTTAAAAAAACCTCCTCTAGAAAA	900
	Db	6881	ATTGATACAAATCGGGAGTGGGCTTCTGCGAAACTTAAAAAAACCTCCTCTAGAAAA	6940
QY		901	ATTCCGAGTGAAGATTGTCTTTTACAGTTTGTATCAAAACGAGGCCAAAAACATCAGTGGT	960
Db		6941	ATTCCGAGTGAAGATTGTCTTTTACAGTTTGTATCAAAACGAGGCCAAAAACATCAGTGGT	7000
	QY	961	CAGAGTCCGGCGGAATCTTCTCCGACCCAGGAGCAACACACACAACTGAGACACACAAA	1020
	Db	7001	CAGAGTCCGGCGGAATCTTCTCCGACCCAGGAGCAACACACACAACTGAGACACACAAA	7060
QY		1021	ATCATGGCTTCAGAAAAATCTCTCTGCAATGGTTTCAAGTGCACAGTCAAGGAAGGAGACT	1080
Db		7061	ATCATGGCTTCAGAAAAATCTCTCTGCAATGGTTTCAAGTGCACAGTCAAGGAAGGAGACT	7120
	QY	1081	GCAGTGTGCAATCTAAACACCCTTGCCAATCTCCAAGTGTCCCAATCCCTCCACAACC	1140
	Db	7121	GCAGTGTGCAATCTAAACACCCTTGCCAATCTCCAAGTGTCCCAATCCCTCCACAACC	7180
QY		1141	AAACCAAGTCCGACACACAGCACCCATATACACCCGTGTATAACTTGACATCTCTGAG	1200
Db		7181	AAACCAAGTCCGACACACAGCACCCATATACACCCGTGTATAACTTGACATCTCTGAG	7240
	QY	1201	GCAACTCAAGTTGAAACAAATCATCCGACAGAACAGACAAACGACAGCAGAGCTCCGACACT	1260
	Db	7241	GCAACTCAAGTTGAAACAAATCATCCGACAGAACAGACAAACGACAGCAGAGCTCCGACACT	7300
QY		1261	CCCTCTGCCACGACCGCAGCCGGAACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC	1320
Db		7301	CCCTCTGCCACGACCGCAGCCGGAACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC	7360
	QY	1321	ACTGACTTCTTGAGACCCCGCACACACAAAGTCCCCAAAAACACAGCAGAGACCGCTGGC	1380
	Db	7361	ACTGACTTCTTGAGACCCCGCACACACAAAGTCCCCAAAAACACAGCAGAGACCGCTGGC	7420
QY		1381	AACACAAACTCATACACCAAGATACCGGAGAGAGAGTCCACAGCAGCGGGAAGCTAGGC	1440
Db		7421	AACACAAACTCATACACCAAGATACCGGAGAGAGAGTCCACAGCAGCGGGAAGCTAGGC	7480
	QY	1441	TTAATTACCAATATCTATTGCTGGAGTCGACGACTGATCAACGCGCGGAGAACTCGA	1500
	Db	7481	TTAATTACCAATATCTATTGCTGGAGTCGACGACTGATCAACGCGCGGAGAACTCGA	7540
QY		1501	CGA 1503	
Db		7541	AGA 7543	

RESULT 4	AR404895	AR404895	2298 bp	DNA	linear	PAT 18-DEC-2003
LOCUS	Sequence 1 from patent US 6630144.					
DEFINITION	AR404895					
ACCESSION	AR404895.1	GI:40153661				
VERSION	Unknown.					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 2298)					
AUTHORS	Hart, M.K., Wilson, J.A. and Schmaljohn, A.L.					
TITLE	Monoclonal antibodies to Ebola glycoprotein					
JOURNAL	Patent: US 6630144-A 1 07-OCT-2003;					
FEATURES	Location/Qualifiers					
source	1..2298					
ORIGIN	/organism="unknown"					
	/mol_type="genomic DNA"					
Query Match	66.6%;	Score 1499.8;	DB 6;	Length 2298;		
Best Local Similarity	99.9%;	Pred. No. 0;				
Matches 1501;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;		
QY	1	ATGGGGGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT	60			
DB	59	ATGGGGGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT	118			
QY	61	CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCAAAT	120			
DB	119	CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCAAAT	178			
QY	121	AGCATTACAGGTTAGTATGATGTCGACAACTAGTTTGTGTCGACAACTGTCTCCACA	180			
DB	179	AGCATTACAGGTTAGTATGATGTCGACAACTAGTTTGTGTCGACAACTGTCTCCACA	238			
QY	181	AATCAATTACAGATCAGTTGGATCTCGAAGGGAATGGAGTGGCACTGACGTGCCA	240			
DB	239	AATCAATTACAGATCAGTTGGATCTCGAAGGGAATGGAGTGGCACTGACGTGCCA	298			
QY	241	TCTGCAATTAAGATGGGGTTCCAGTCCGGTGTCCCAACCAAGGTGGTCAATTATGAA	300			
DB	299	TCTGCAATTAAGATGGGGTTCCAGTCCGGTGTCCCAACCAAGGTGGTCAATTATGAA	358			
QY	301	GCTGGTGAATGGGCTGAAACTGCTACATCTTGAATCAAAAACCTGACGGGATGAG	360			
DB	359	GCTGGTGAATGGGCTGAAACTGCTACATCTTGAATCAAAAACCTGACGGGATGAG	418			
QY	361	TGTTCTACAGCAGCGCAGACGGGATTCCGGGCTTCCCGGTTCCCGGTTATGTGCACAA	420			
DB	419	TGTTCTACAGCAGCGCAGACGGGATTCCGGGCTTCCCGGTTCCCGGTTATGTGCACAA	478			
QY	421	GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATTAAGAGGGTCTTTCTTC	480			
DB	479	GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATTAAGAGGGTCTTTCTTC	538			
QY	481	CTGTATGATCGATCTTCTTCCAGTATCTACCGAGGAACGACTTTCCTGAAGGTGTC	540			
DB	539	CTGTATGATCGATCTTCTTCCAGTATCTACCGAGGAACGACTTTCCTGAAGGTGTC	598			
QY	541	GTTCGATTTCTGATCTGCCCCAGCTAAGAGGACTTCTTCAGCTCAGACCCCTTGAGA	600			
DB	599	GTTCGATTTCTGATCTGCCCCAGCTAAGAGGACTTCTTCAGCTCAGACCCCTTGAGA	658			
QY	601	GAGCCGGTCAATGCAACCGAGGACCGTCTAGTGGCTACTATTCTTACCAATATAGATAT	660			
DB	659	GAGCCGGTCAATGCAACCGAGGACCGTCTAGTGGCTACTATTCTTACCAATATAGATAT	718			
QY	661	CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTCGAGGTTGACAAATTGACC	720			
DB	719	CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTCGAGGTTGACAAATTGACC	778			
QY	721	TACGTCCAACCTTGAATCAAGATTCAACCAAGTTTCTGCTCCAGCTGAATGAGACAATA	780			
DB	779	TACGTCCAACCTTGAATCAAGATTCAACCAAGTTTCTGCTCCAGCTGAATGAGACAATA	838			
QY	781	TATACAACTGGGAAAAGAGCAATACACCGGAAAACCTAATTTGGAAGGTCAACCCCGAA	840			
DB	839	TATACAACTGGGAAAAGAGCAATACACCGGAAAACCTAATTTGGAAGGTCAACCCCGAA	898			
QY	841	ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACTTAAAAAAAACCCTCACTAGAAAA	900			
DB	899	ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACTTAAAAAAAACCCTCACTAGAAAA	958			
QY	901	ATTGCACTGAAGAGTTGCTTTTCAAGTTGATCAAAACGGAGCCAAAACATCATGTGT	960			
DB	959	ATTGCACTGAAGAGTTGCTTTTCAAGTTGATCAAAACGGAGCCAAAACATCATGTGT	1018			
QY	961	CAGAGTCCGGGCGGAACTTCTTCGACCCAGGAGCAACACAACTGAGAGACACAAA	1020			
DB	1019	CAGAGTCCGGGCGGAACTTCTTCGACCCAGGAGCAACACAACTGAGAGACACAAA	1078			
QY	1021	ATCATGGCTTCAGAAAAATTCCTCTCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT	1080			
DB	1079	ATCATGGCTTCAGAAAAATTCCTCTCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT	1138			
QY	1081	GCAGTGTGCGCATCTAAACAAACCTTGGCCCAATCTCCACGAGTCCCAATCCCTCAACCC	1140			
DB	1139	GCAGTGTGCGCATCTAAACAAACCTTGGCCCAATCTCCACGAGTCCCAATCCCTCAACCC	1198			
QY	1141	AAACAGAGTCCGGCAACAGACCCCATTAATACACCCGTGTATATACTTGACATCTCTGAG	1200			
DB	1199	AAACAGAGTCCGGCAACAGACCCCATTAATACACCCGTGTATATACTTGACATCTCTGAG	1258			
QY	1201	GCAACTCAAGTTGAACAACTCACCAGCAACAGACAAACAGACAGCAGCTCCGACACT	1260			
DB	1259	GCAACTCAAGTTGAACAACTCACCAGCAACAGACAAACAGACAGCAGCTCCGACACT	1318			
QY	1261	CCCTCTGCGACGCGCAGCGGACCCCAAAAGCAGAGAACACCAACACGAGCAAGAGC	1320			
DB	1319	CCCTCTGCGACGCGCAGCGGACCCCAAAAGCAGAGAACACCAACACGAGCAAGAGC	1378			
QY	1321	ACTGACTTCTGGAGCCCGCCACCAACAGATGTCCTCCCAAAACACAGCGAGACCGTGGC	1380			
DB	1379	ACTGACTTCTGGAGCCCGCCACCAACAGATGTCCTCCCAAAACACAGCGAGACCGTGGC	1438			
QY	1381	AACAAACACTCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGGAGGCTAGGC	1440			
DB	1439	AACAAACACTCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGGAGGCTAGGC	1498			
QY	1441	TTAATTACCAATATCTTATTTGCTGGAGTCGAGACTGTATCAGCGGGGAGAAAGACTCGA	1500			
DB	1499	TTAATTACCAATATCTTATTTGCTGGAGTCGAGACTGTATCAGCGGGGAGAAAGACTCGA	1558			
QY	1501	CGA 1503				
DB	1559	AGA 1561				
RESULT 5	AX092102		2298 bp	DNA	linear	PAT 29-AUG-2002
LOCUS	Sequence 1 from Patent WO0116183.					
DEFINITION	AX092102					
ACCESSION	AX092102					
VERSION	AX092102.1	GI:13444354				
KEYWORDS	Zaire ebolavirus (ZBOV)					
SOURCE	Zaire ebolavirus					
ORGANISM	Zaire ebolavirus					
REFERENCE	1					
AUTHORS	Hart, M.K., Wilson, J.A. and Schmaljohn, A.L.					
TITLE	Monoclonal antibodies and vaccines against epitopes on the ebola virus glycoprotein					

JOURNAL Patent: WO 0116193-A 1 08-MAR-2001;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES Location/Qualifiers
source 1..2298
/organism="Zaire ebolavirus"
/mol_type="unassigned DNA"
/db_xref="taxon:186538"

ORIGIN

Query Match 66.6%; Score 1499.8; DB 6; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGGCGGTTACAGGAATATTGCAAGTACCTGTTGATCGAATCAAGAGGACATCATCTTT 60
DB 59 ATGGCGGTTACAGGAATATTGCAAGTACCTGTTGATCGAATCAAGAGGACATCATCTTT 118
QY 61 CTTTGGGTAATTTATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAACCAAT 120
DB 119 CTTTGGGTAATTTATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAACCAAT 178
QY 121 AGCATTATACAGGTTAGTGTGACAAACTAGTTTGTGTCGACAAACTGTTCATCCACA 180
DB 179 AGCATTATACAGGTTAGTGTGACAAACTAGTTTGTGTCGACAAACTGTTCATCCACA 238
QY 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
DB 239 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 298
QY 241 TCTGCAACTAAAGATGGGCTTCAGGTCGGGTGCCCAACAAAGTGCTCAATTTATGAA 300
DB 299 TCTGCAACTAAAGATGGGCTTCAGGTCGGGTGCCCAACAAAGTGCTCAATTTATGAA 358
QY 301 GCTGTGAATGGGCTGAAAACTGCTTCAATCTTGAATCAAAATCAAAACCTGACGGAGTGAG 360
DB 359 GCTGTGAATGGGCTGAAAACTGCTTCAATCTTGAATCAAAACCTGACGGAGTGAG 418
QY 361 TGTCTACAGACGCCACGACGAGATTCGGGCTTCCCGGTCGGGTGCGGTATGTGACAAA 420
DB 419 TGTCTACAGACGCCACGACGAGATTCGGGCTTCCCGGTCGGGTGCGGTATGTGACAAA 478
QY 421 GTATCAGGAACGGGACCGGTGTCGGGACACTTTCCTTCCATAAGAGGGTCTTCTTC 480
DB 479 GTATCAGGAACGGGACCGGTGTCGGGACACTTTCCTTCCATAAGAGGGTCTTCTTC 538
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAAGACTTTCGCTGAAGGTGTC 540
DB 539 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAAGACTTTCGCTGAAGGTGTC 598
QY 541 GTTGCATTTCTGATATCTGCCCAAGCTAAGAGGACTTCTTCAGCTCACAACCCCTTGAGA 600
DB 599 GTTGCATTTCTGATATCTGCCCAAGCTAAGAGGACTTCTTCAGCTCACAACCCCTTGAGA 658
QY 601 GAGCCGGTCAATGCAACGGAGACCGGTCTAGTGGCTACTATTCTACCAATATAGATAT 660
DB 659 GAGCCGGTCAATGCAACGGAGACCGGTCTAGTGGCTACTATTCTACCAATATAGATAT 718
QY 661 CAGGCTACCGGTTTGGAACTAAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
DB 719 CAGGCTACCGGTTTGGAACTAAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 778
QY 721 TACGTCCTCAACTTGAATCAAGATTTCAACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
DB 779 TACGTCCTCAACTTGAATCAAGATTTCAACACAGTTTCTGCTCCAGCTGAATGAGACAATA 838
QY 781 TATACAGTGGGAAAGGAGCAATACCAAGGAAACTAATTTGGAAGTCAACCCGCA 840
DB 839 TATACAGTGGGAAAGGAGCAATACCAAGGAAACTAATTTGGAAGTCAACCCGCA 898
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAAAAAACCCTCACTAGAAAA 900
DB 899 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAAAAAACCCTCACTAGAAAA 958
```

```
QY 901 ATTGCGAGTGAAGAGTTGTCTTTCACAGTTGTATCAAAACGAGCCAAAACATCAGTGT 960
DB 959 ATTGCGAGTGAAGAGTTGTCTTTCACAGTTGTATCAAAACGAGCCAAAACATCAGTGT 1018
QY 961 CAGAGTCCGGCGGAACTTCTTCGACCCAGGGACCAACAACTCTGAAGACCAAAA 1020
DB 1019 CAGAGTCCGGCGGAACTTCTTCGACCCAGGGACCAACAACTCTGAAGACCAAAA 1078
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
DB 1079 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1138
QY 1081 GCAGTGTGCGATCTAAACAACTTTCGACCAATCTCCAGAGTCCCAATCCCTCACAACC 1140
DB 1139 GCAGTGTGCGATCTAAACAACTTTCGACCAATCTCCAGAGTCCCAATCCCTCACAACC 1198
QY 1141 AAACAGGTCGGGACAAACAGCACCACCAATAATACACCGTGTATAAACTTGACATCTCTGAG 1200
DB 1199 AAACAGGTCGGGACAAACAGCACCACCAATANTACACCGTGTATAAACTTGACATCTCTGAG 1258
QY 1201 GCAACTCAAGTTGAACCAATCAACCGCAGAACAGACAAACGACAGCAGCCTCCGACACT 1260
DB 1259 GCAACTCAAGTTGAACCAATCAACCGCAGAACAGACAAACGACAGCAGCCTCCGACACT 1318
QY 1261 CCCTCTGCCAGCAGCGGACCGGACCCCAAAAGCAGAGAACCAACACGAGCAAGGC 1320
DB 1319 CCCTCTGCCAGCAGCGGACCGGACCCCAAAAGCAGAGAACCAACACGAGCAAGGC 1378
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACACAGCAGCAGCCTCGGC 1380
DB 1379 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACACAGCAGCAGCCTCGGC 1438
QY 1381 AACAAACAACTCATATCAACCAAGATACCGGAGAAAGAGTGCCAGACGCGGAAGCTAGGC 1440
DB 1439 AACAAACAACTCATATCAACCAAGATACCGGAGAAAGAGTGCCAGACGCGGAAGCTAGGC 1498
QY 1441 TTAATTACCAATATCTTGTGGAGTGCAGGACTGATCAAGCGGGGAGAGAACTCGA 1500
DB 1499 TTAATTACCAATATCTTGTGGAGTGCAGGACTGATCAAGCGGGGAGAGAACTCGA 1558
QY 1501 CGA 1503
DB 1559 AGA 1561
```

RESULT 6
EVU31033
LOCUS
DEFINITION Zaire Ebola virus envelope glycoprotein (GP) gene, complete cds,
secreted small glycoprotein (sgp) gene, complete cds.
ACCESSION U31033 S59050
VERSION U31033.1 GI:1141778
KEYWORDS
SOURCE Zaire ebolavirus (ZEBOV)
ORGANISM Zaire ebolavirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE 1 (sites)
AUTHORS Volchkov,V.E., Becker,S., Volchkova,V.A., Ternovoj,V.A.,
Kotov,A.N., Netesov,S.V. and Klenk,H.D.
GP mRNA of Ebola virus is edited by the Ebola virus polymerase and
by T7 and vaccinia virus polymerases
Virology 214 (2), 421-430 (1995)
JOURNAL MEDLINE 96130181
PUBMED 8553543
REFERENCE 2 (sites)
AUTHORS Volchkov,V.E., Blinov,V.M., Kotov,A., Chepurinov,A. and Netesov,S.
TITLE The full-length nucleotide sequence of the Ebola virus
JOURNAL Thesis (1993)
REFERENCE 3 (sites)
AUTHORS Volchkov,V.E., Blinov,V.M. and Netesov,S.V.
TITLE The envelope glycoprotein of Ebola virus contains an
immunosuppressive-like domain similar to oncogenic retroviruses

JOURNAL	FEBS Lett. 305 (3), 181-184 (1992)	Best Local Similarity	99.9%;	Pred. No. 0;						
MEDLINE	93231342	Matches 1501;	Conservative	0;	Mismatches	1;	Indels	1;	Gaps	
PUBMED	1299611									
REFERENCE	4 (bases 1 to 2406)									
AUTHORS	Volchkov, V.E.									
TITLE	Direct Submission									
JOURNAL	Submitted (06-JUL-1995) Viktor E. Volchkov, Institut fuer Virologie, Philipps-Universitaet, Robert-Koch-Str.17, Marburg 35037, Germany									
COMMENT	On Dec 31, 1995 this sequence version replaced gi:299674.									
FEATURES	Location/Qualifiers									
source	1. .2406									
	/organism="Zaire ebolavirus"									
	/mol_type="genomic RNA"									
	/strain="Zaire"									
	/db_xref="taxon:186538"									
CDS	join(140..1162,1162..2169)									
	/note="Synthesis of the full length EBO GP occurs from edited form of GP mRNA; one additional A residue inserts in EBO GP editing site."									
	/citation={3}									
	/citation={1}									
	/citation={2}									
	/codon_start=1									
	/evidence=experimental									
	/product="envelope glycoprotein"									
	/protein_id="AAA96744.1"									
	/db_xref="GI:1141779"									
	/translation="MGVIGILQLPRDRFKRTSFFLWVILFORTFSILPLGVHNSTLQ VSDVKLVCRDKLSLNQURSVGLNLENGVAIDVPSATKRWGFRSGVPPKVVNTEAG EWAENCYNLEIKKDPGSECLPAAPDGIRGFPFRXYHKVSGTGPCAGDFAFHKEGAF LVDRLASTVIYRGTFPAEGVAFLLIPQAKKDFSSHPLEPNVATEDPSSGYSTTI RYQATGFGNTEYLFVDVNLTVOLESRFTQFLQLNETIYTSGKRSNTTGLKIWK VNPEDITTTGEAFWETKTSLEKAVKSCLSQLYQTEPKTSVWVRRELLPTQGPQT QUKTKMASENSAMVQHSQGREAAVSHLTLLAIISTSPQSLLTKGPDNSHNTP VYKLDISEATQEQHRRITDNDSTASDTPSATTAAGPPKAENTNKSDFLDPAITT SPQHSNTEAGNNTTHODTEESASSGKLGLITNAGVAGLITGGRTRREAI VNAQ PCKNPLHWYTTQDEGAALGLAWIPYFGPAEGIVIEGLMHNQDGLICGLRLANETT QALQLFLRATTFRTFSILNRKAIDFLQRWGTCHILGPDCCEIHPDWTNKTIDKID QIHHFVDKTLDPQGDNDNWTGWQWIPAGIVGTGVIIAVIALFCICKFVF"									
gene	140..1234									
	/gene="SGP"									
CDS	140..1234									
	/gene="sgp"									
	/note="synthesis on non-edited form of EBO GP mRNA"									
	/species=" "									
	/citation={2}									
	/codon_start=1									
	/evidence=experimental									
	/product="secreted small glycoprotein"									
	/protein_id="AAA96745.1"									
	/db_xref="GI:1141780"									
	/translation="MGVIGILQLPRDRFKRTSFFLWVILFORTFSILPLGVHNSTLQ VSDVKLVCRDKLSLNQURSVGLNLENGVAIDVPSATKRWGFRSGVPPKVVNTEAG EWAENCYNLEIKKDPGSECLPAAPDGIRGFPFRXYHKVSGTGPCAGDFAFHKEGAF LVDRLASTVIYRGTFPAEGVAFLLIPQAKKDFSSHPLEPNVATEDPSSGYSTTI RYQATGFGNTEYLFVDVNLTVOLESRFTQFLQLNETIYTSGKRSNTTGLKIWK VNPEDITTTGEAFWETKTSLEKAVKSCLSQLYQTEPKTSVWVRRELLPTQGPQT QUKTKSWLQKTLPLQWFKCTVKEGLQCR1"									
misc_feature	1019..1025									
	/gene="sgp"									
	/note="RNA editing site; synthesis of full length GP proceeds via RNA editing; one additional A residue inserts during transcription of the Ebola virus GP mRNA in region; contains 7 consecutive A residues."									
	/citation={1}									
	/citation={2}									
	/evidence=experimental									
polya_signal	2397..2406									
	/evidence=experimental									
ORIGIN										
Query Match	66.1%; Score 1489.4; DB 14; Length 2406;									

```

Db 1159 ATCATGGCTTCAGAAAATTCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1218
QY 1081 GCAGTGTGCATCTAAACACCTTCGCCAATCTCCACAGATGCCCAATCCCTCACAACC 1140
Db 1219 GCAGTGTGCATCTAAACACCTTCGCCAATCTCCACAGATGCCCAATCCCTCACAACC 1278
QY 1141 AAACAGGTCCGGACACAGACCCATTAATACACCGGTGTATAAATCTTGACATCTCTGAG 1200
Db 1279 AAACAGGTCCGGACACAGACCCATTAATACACCGGTGTATAAATCTTGACATCTCTGAG 1338
QY 1201 GCAACTCAAGTTGAACAACATCACCAGCAGAGACAGACAACAGCAGACAGCCTCCGACACT 1260
Db 1339 GCAACTCAAGTTGAACAACATCACCAGCAGAGACAGACAACAGCAGACAGCCTCCGACACT 1398
QY 1261 CCCTCTGCACACACCGAGCGGACCCCAAAAGCAGAGACACCAACAGCAGCAGAGC 1320
Db 1399 CCCTCTGCACACACCGAGCGGACCCCAAAAGCAGAGACACCAACAGCAGCAGAGC 1458
QY 1321 ACTGACTTCTTGGACCCCGCCACACCAACAGTCCCAAAACACAGCAGAGACCGCTGCG 1380
Db 1459 ACTGACTTCTTGGACCCCGCCACACCAACAGTCCCAAAACACAGCAGAGACCGCTGCG 1518
QY 1381 AACAAACACACTCATCACCAGATACCGGAGAGAGTGCAGCAGCGGGAAGCTAGGC 1440
Db 1519 AACAAACACACTCATCACCAGATACCGGAGAGAGTGCAGCAGCGGGAAGCTAGGC 1578
QY 1441 TTAATTACAACTATTGCTGGAGTCGAGACTGATCAGAGCGGAGAGAACTCGA 1500
Db 1579 TTAATTACAACTATTGCTGGAGTCGAGACTGATCAGAGCGGAGAGAACTCGA 1638
QY 1501 CGA 1503
Db 1639 AGA 1641

RESULT 7
EVU23187
LOCUS Zaire Ebola virus Mayinga strain glycoprotein (GP) gene, complete
DEFINITION
cde.
ACCESSION U23187
VERSION U23187.1 GI:1041204
KEYWORDS
SOURCE Zaire ebolavirus (ZEBOV)
ORGANISM Zaire ebolavirus
VIRUSES; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE
1 (bases 1 to 2408)
Sanchez,A., Trappier,S.G., Mahy,B.W., Peters,C.J. and Nichol,S.T.
The virion glycoproteins of Ebola viruses are encoded in two
reading frames and are expressed through transcriptional editing
Proc. Natl. Acad. Sci. U.S.A. 93 (8), 3602-3607 (1996)
96195018
PUBMED 8622982
REFERENCE
2 (bases 1 to 2408)
Sanchez,A., Trappier,S., Conaty,A.L., Brammer,L., Mahy,B.J.W.,
Peters,C.J. and Nichol,S.T.
Direct Submission
Submitted (22-MAR-1995) Anthony Sanchez, Special Pathogens Branch,
Division of Viral and Rickettsial Diseases, Centers for Disease
Control and Prevention, 1600 Clifton Road, Bldg. 15, Room SB611,
Mail Stop G14, Atlanta, GA 30333, USA
LOCATION/Qualifiers
1. .2408
/organism="Zaire ebolavirus"
/mol_type="genomic RNA"
/strain="Mayinga 1976"
/db_xref="taxon:186538"
/ncbi="subtype: Zaire"
1. .141
1. .14
142. .2171
5'UTR
misc_signal
note="transcription start site"
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616
2617
2618
2619
2620
2621
2622
2623
2624
2625
2626
2627
2628
2629
2630
2631
2632
2633
2634
2635
2636
2637
2638
2639
2640
2641
2642
2643
2644
2645
2646
2647
2648
2649
2650
2651
2652
2653
2654
2655
2656
2657
2658
2659
2660
2661
2662
2663
2664
2665
2666
2667
2668
2669
2670
2671
2672
2673
2674
2675
2676
2677
2678
2679
2680
2681
2682
2683
2684
2685
2686
2687
2688
2689
2690
2691
2692
2693
2694
2695
2696
2697
2698
2699
2700
2701
2702
2703
2704
2705
2706
2707
2708
2709
2710
2711
2712
2713
2714
2715
2716
2717
2718
2719
2720
2721
2722
2723
2724
2725
2726
2727
2728
2729
2730
2731
2732
2733
2734
2735
2736
2737
2738
2739
2740
2741
2742
2743
2744
2745
2746
2747
2748
2749
2750
2751
2752
2753
2754
2755
2756
2757
2758
2759
2760
2761
2762
2763
2764
2765
2766
2767
2768
2769
2770
2771
2772
2773
2774
2775
2776
2777
2778
2779
2780
2781
2782
2783
2784
2785
2786
2787
2788
2789
2790
2791
2792
2793
2794
2795
2796
2797
2798
2799
2800
2801
2802
2803
2804
2805
2806
2807
2808
2809
2810
2811
2812
2813
2814
2815
2816
2817
2818
2819
2820
2821
2822
2823
2824
2825
2826
2827
2828
2829
2830
2831
2832
2833
2834
2835
2836
2837
2838
2839
2840
2841
2842
2843
2844
2845
2846
2847
2848
2849
2850
2851
2852
2853
2854
2855
2856
2857
2858
2859
2860
2861
2862
2863
2864
2865
2866
2867
2868
2869
2870
2871
2872
2873
2874
2875
2876
2877
2878
2879
2880
2881
2882
2883
2884
2885
2886
2887
2888
2889
2890
2891
2892
2893
2894
2895
2896
2897
2898
2899
2900
2901
2902
2903
2904
2905
2906
2907
2908
2909
2910
2911
2912
2913
2914
2915
2916
2917
2918
2919
2920
2921
2922
2923
2924
2925
2926
2927
2928
2929
2930
2931
2932
2933
2934
2935
2936
2937
2938
2939
2940
2941
2942
2943
2944
2945
2946
2947
2948
2949
2950
2951
2952
2953
2954
2955
2956
2957
2958
2959
2960
2961
2962
2963
2964
2965
2966
2967
2968
2969
2970
2971
2972
2973
2974
2975
2976
2977
2978
2979
2980
2981
2982
2983
2984
2985
2986
2987
2988
2989
2990
2991
2992
2993
2994
2995
2996
2997
2998
2999
3000
3001
3002
3003
3004
3005
3006
3007
3008
3009
3010
3011
3012
3013
3014
3015
3016
3017
3018
3019
3020
3021
3022
3023
3024
3025
3026
3027
3028
3029
3030
3031
3032
3033
3034
3035
3036
3037
3038
3039
3040
3041
3042
3043
3044
3045
3046
3047
3048
3049
3050
3051
3052
3053
3054
3055
3056
3057
3058
3059
3060
3061
3062
3063
3064
3065
3066
3067
3068
3069
3070
3071
3072
3073
3074
3075
3076
3077
3078
3079
3080
3081
3082
3083
3084
3085
3086
3087
3088
3089
3090
3091
3092
3093
3094
3095
3096
3097
3098
3099
3100
3101
3102
3103
3104
3105
3106
3107
3108
3109
3110
3111
3112
3113
3114
3115
3116
3117
3118
3119
3120
3121
3122
3123
3124
3125
3126
3127
3128
3129
3130
3131
3132
3133
3134
3135
3136
3137
3138
3139
3140
3141
3142
3143
3144
3145
3146
3147
3148
3149
3150
3151
3152
3153
3154
3155
3156
3157
3158
3159
3160
3161
3162
3163
3164
3165
3166
3167
316
```

QY 301 GCTGTGAATGGGCTGAAACCTGCTACAATCTTGAAATCAAAAAACCTGACGGGAGTGAG 360
Db 442 GCTGTGAATGGGCTGAAACCTGCTACAATCTTGAAATCAAAAAACCTGACGGGAGTGAG 501
QY 361 TGTCTACAGCAGCGCAGACGGGATTCGGGGCTTCCCGCGTGCCTGATGTGCACAAA 420
Db 502 TGTCTACAGCAGCGCAGACGGGATTCGGGGCTTCCCGCGTGCCTGATGTGCACAAA 561
QY 421 GTATCAGGAACGGGACCGGTGTCGGAGACTTTGCTTCCATAAAGAGGGTCTTCTTC 480
Db 562 GTATCAGGAACGGGACCGGTGTCGGAGACTTTGCTTCCATAAAGAGGGTCTTCTTC 621
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
Db 622 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 681
QY 541 GTTGCATTTCTGATACTGCCCCAAGCTAAGAAGGACTTCTTACGCTCACACCCCTTGAGA 600
Db 682 GTTGCATTTCTGATACTGCCCCAAGCTAAGAAGGACTTCTTACGCTCACACCCCTTGAGA 741
QY 601 GAGCCGGTCAATGCAACGGGAGCAGCCGCTAGTGGCTACTATTCTACCAATATTAGATAT 660
Db 742 GAGCCGGTCAATGCAACGGGAGCAGCCGCTAGTGGCTACTATTCTACCAATATTAGATAT 801
QY 661 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
Db 802 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 861
QY 721 TAGCTCCAACTTGAATCAAGATTTCACACAGTTTCTGCTCCAGTGAATGAGACAATA 780
Db 862 TAGCTCCAACTTGAATCAAGATTTCACACAGTTTCTGCTCCAGTGAATGAGACAATA 921
QY 781 TATACAAGTGGGAAAGGAGCAATACCGGGAAACTAATTTGGAAGTCAACCCCGAA 840
Db 922 TATACAAGTGGGAAAGGAGCAATACCGGGAAACTAATTTGGAAGTCAACCCCGAA 981
QY 841 ATTGATACAACAAATCGGGAGTGGGCTTCTGGGAACTAAAAAAACCTCTACTAGAAA 900
Db 982 ATTGATACAACAAATCGGGAGTGGGCTTCTGGGAACT-AAAAAAACCTCTACTAGAAA 1040
QY 901 ATTGAGTGAAGAGTTGTCTTTCAGTTGATATCAACGGAGCCAAAACATCATGTTGT 960
Db 1041 ATTGAGTGAAGAGTTGTCTTTCAGTTGATATCAACGGAGCCAAAACATCATGTTGT 1100
QY 961 CAGAGTCCGGCGGCACTTCTTCGACCCAGGAGCAACACAACTGCAACACACACAAA 1020
Db 1101 CAGAGTCCGGCGGCACTTCTTCGACCCAGGAGCAACACAACTGCAACACACACAAA 1160
QY 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1080
Db 1161 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1220
QY 1081 GAGTGTGCGATCTAAACAACCTTGCACAATCTCCAGAGTCCCAATCCCTCTCAACCC 1140
Db 1221 GAGTGTGCGATCTAAACAACCTTGCACAATCTCCAGAGTCCCAATCCCTCTCAACCC 1280
QY 1141 AAACCAAGTCCGAGCAACAGCAACCAATATACACCGTGTATATAAATCTTCAATCTCTGAG 1200
Db 1281 AAACCAAGTCCGAGCAACAGCAACCAATATACACCGTGTATATAAATCTTCAATCTCTGAG 1340
QY 1201 GCAACTCAAGTTGAAACAACATCACCGCAGACAGCAACGACAGCAACGACGAGAGAGC 1260
Db 1341 GCAACTCAAGTTGAAACAACATCACCGCAGACAGCAACGACAGCAACGACGAGAGAGC 1400
QY 1261 CCTCTGCCACGACCGGAGCGGACCCCAAAAGCAGAGAAACCAACGACGAGAGAGC 1320
Db 1401 CCTCTGCCACGACCGGAGCGGACCCCAAAAGCAGAGAAACCAACGACGAGAGAGC 1460
QY 1321 ACTGACTTCTGTGACCCCGCACCAACAAAGTCCCAAAACCAACGACGAGAGCGCTGCG 1380
Db 1461 ACTGACTTCTGTGACCCCGCACCAACAAAGTCCCAAAACCAACGACGAGAGCGCTGCG 1520
QY 1381 AACAAACAACATCTATCAACCAAGATACCGGAGAGAGAGTGCCAGCAGCGGGAAGCTTAGC 1440

Db 1521 AACAAACAACATCTATCAACCAAGATACCGGAGAGAGTGCCAGCAGCGGGAAGCTAGGC 1580
QY 1441 TTAATTACCAATACTATTGCTGGAGTCCGAGACTGTATCAGCGCGGAGAACTCGA 1500
Db 1581 TTAATTACCAATACTATTGCTGGAGTCCGAGACTGTATCAGCGCGGAGAACTCGA 1640
QY 1501 CGA 1503
Db 1641 AGA 1643
RESULT 8
AF086833
LOCUS Zaire Ebola virus strain Mayinga, complete genome. VRL 28-AUG-2002
DEFINITION Zaire Ebola virus strain Mayinga, complete genome.
ACCESSION AF086833
VERSION AF086833.2 GI:10141003
KEYWORDS
SOURCE Zaire ebolavirus (ZEBOV)
ORGANISM Zaire ebolavirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE 1 (bases 1 to 18959)
AUTHORS Bukreyev,A.A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.
TITLE The VP35 and VP40 proteins of filoviruses. Homology between Marburg and Ebola viruses
JOURNAL FEBS Lett. 322 (1), 41-46 (1993)
MEDLINE 93245956
PUBMED 8482365
REFERENCE 2 (bases 1 to 18959)
AUTHORS Volchkov,V.E., Becker,S., Volchkova,V.A., Ternovoj,V.A.,
Kotov,A.N., Netesov,S.V. and Klenk,H.D.
TITLE GP mRNA of Ebola virus is edited by the Ebola virus polymerase and by T7 and vaccinia virus polymerases
JOURNAL Virology 214 (2), 421-430 (1995)
MEDLINE 96130181
PUBMED 8553543
REFERENCE 3 (bases 1 to 18959)
AUTHORS Volchkov,V.E., Feldmann,H., Volchkova,V.A. and Klenk,H.D.
TITLE Processing of the Ebola virus glycoprotein by the proprotein convertase furin
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (10), 5762-5767 (1998)
MEDLINE 98245155
PUBMED 9576958
REFERENCE 4 (bases 1 to 18959)
AUTHORS Volchkov,V.E., Volchkova,V.A., Slenczka,W., Klenk,H.D. and Feldmann,H.
TITLE Release of viral glycoproteins during Ebola virus infection
JOURNAL Virology 245 (1), 110-119 (1998)
MEDLINE 98277077
PUBMED 9614872
REFERENCE 5 (bases 1 to 18959)
AUTHORS Volchkov,V.E., Volchkova,V.A., Chepurinov,A.A., Blinov,V.M., Dolnik,O., Netesov,S.V. and Feldmann,H.
TITLE Characterization of the L gene and 5' trailer region of Ebola virus
JOURNAL J. Gen. Virol. 80 (Pt 2), 355-362 (1999)
MEDLINE 99171703
PUBMED 10073695
REFERENCE 6 (bases 1 to 18959)
AUTHORS Volchkov,V.E.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-1998) Institute of Virology, Philipps-University Marburg, Robert-Koch-Str. 17, Marburg 35037, Germany
PUBMED 10073695
REFERENCE 7 (bases 1 to 18959)
AUTHORS Volchkov,V.E.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-2000) Institute of Virology, Philipps-University Marburg, Robert-Koch-Str. 17, Marburg 35037, Germany
REMARK Sequence update by submitter
COMMENT On Sep 15, 2000 this sequence version replaced gi:4262346.
FEATURES
location/Qualifiers
1..18959
source

```

/organism="Zaire ebolavirus"
/mol_type="genomic RNA"
/strain="Mayinga"
/db_xref="taxon:186538"
1..55
/citation=[5]
/function="regulation or initiation of RNA replication"
56..3026
/gene="NP"
56..3026
/gene="NP"
56..3026
/product="nucleoprotein"
56..67
/gene="NP"
/feature="putative; transcription start signal"
/citation=[5]
470..2689
/gene="NP"
/function="encapsidation of genomic RNA"
/codon_start=1
/product="nucleoprotein"
/protein_id="AADI4590.1"
/db_xref="GI:4262355"
/translation="WDSRPQKIWAPSLTESDMYHKILTAGLSVQOGIVQRVIPVY
QYNLEECQLIIQAFEGVDFQSDSFLMLCLHAYQGYKLFLESGAVKYLEGH
GRFEVKRDRGKLEELLPAVSSGKNKIRTLAAMPBEEETTEANAGQFSPASFLPK
LVGEKACLEKVRQIQVHAEQGLIQTPTAQSVGHMMVI FRLMRTNFIKFLLIHQG
MHMVAGHDANDVINSVAQARFGLIIVKTVLDHILOKTERGVRLHPARTAKVNE
VNSKALSSLAKEGVEAPFARLLNLSGVNLEHGLRPQLSAIAGVATAGHSIQQTN
VNSRQVQQLREAEKQLQVAAERELHGLDDQEKILMNFPHQKNEHSFQQTN
AMVIRKELAKITEAITAASPKTSGHYDDDDIPPPGINDDNPCHQDDPDSQ
DTTIPDVVDDPDSGYQSYSENGMNPDDLVLFDLDEDDTKVPVNRSTGGQQ
KNSQKQIEGRQTSRQIQNVPGPRTIHASAPLTDNDRNFPSSGTSRMLTPTIN
EADPLDDADDLSLPQSDDEQDRDGTNRPTVAPVAPVYRSEKELPQDE
QODQDHTOAEARNQSDNTQSEHSPEMYRHILRSQDPDAVLYTHMMKDEPVFVSTSD
GKEYTPPSLEEYPPWLTEKAMNEENRFTVLDGQQPVPVMMHKKFMAILQHQ"
3015..3026
/gene="NP"
3027..3031
/feature="intergenic region"
3032..4407
/gene="VP35"
3032..4407
/gene="VP35"
/product="VP35"
/citation=[1]
3032..3043
/gene="VP35"
/feature="putative; transcription start signal"
/citation=[1]
3129..4151
/gene="VP35"
/function="polymerase complex protein"
/citation=[1]
/codon_start=1
/product="VP35"
/protein_id="AADI4582.1"
/db_xref="GI:4262347"
/translation="MTTIRKRGHTAANTQDRMPGPSPGSHISEQLMTGRIPVSDIF
CDIENPGLCVASQCKQPKPNPKTNSQTQDPCINHSFEEVQVQLASLVQQQTI
ASBSEQRITSLNGLKPVYDMAKTISLNRVCAEMVAKYDLLVMTTGRATATAAATE
AYWAEHGQPPPGPSIYESAIRGKIESRDETVPQSVREAFNNLSNTSTTEENFGKPD
ISAKDLRIMYDHLPGFGTAPHOLVQVICKLGSNSLDIITHAEFOASLAEQSDPOCA
LIQITKRVPIQDAPPPVHISRSDIPRACQKSLRPVPPSPKIDRGWVCVQLQDGK
TLGLKI"
4390..5894
/gene="VP40"
4390..5894
/gene="VP40"
/product="VP40"
/citation=[1]
4390..4401

```

```

/polyA_signal
/feature="transcription start signal"
/citation=[1]
4397..4407
/gene="VP35"
/citation=[1]
4479..5459
/gene="VP40"
/citation=[1]
/codon_start=1
/product="matrix protein"
/protein_id="AADI4583.1"
/db_xref="GI:4262348"
/translation="MRRVILPTAPPEYMEALYVRSNSTIARGNSNTGFLTPEPSVNG
DTPSNLRPIADDTIDHASHTPGSVSSAFLEAMVNVISGPKVLMKQIPIWPLGVAD
QKTSYFSDSTTAALMLASYTTHFGKATNPLVRNRLGFGIPDHLRLLRIGNQPLQLE
FVLPPVLPQVYFTFDLTKLITQPLPAATWDDTPTGSGALRPGISFHKLRPIL
PNKSGKGNAGDLTSPKIQAIMTSLQDKVPIDPTKNIMGIEVPELTVHKLGTGKV
TSKNGQPIIPVLLPKYIGLDPVAPGDLTMTITQDCDTCSPASLPAVIEK"
5883..5894
/gene="VP40"
/citation=[1]
5895..5899
/feature="intergenic region"
5900..8305
/gene="GP"
5900..8305
/gene="GP"
/product="sGP"
/feature="unedited mRNA"
/citation=[2]
5900..5911
/gene="GP"
/feature="putative; transcription start signal"
/citation=[2]
join(6039..6923,6923..8068)
/gene="GP"
/function="receptor binding and fusion"
/feature="an addition A residue is inserted during
transcription; encodes two disulfide linked subunits GP1
and GP2"
/citation=[3]
/citation=[3]
/citation=[4]
/codon_start=1
/exception="RNA editing"
/product="viral spike glycoprotein precursor"
/protein_id="AADI4585.1"
/db_xref="GI:4262350"
/translation="MGVTGILQLPRDRFKRTSPFLMWILFQRTFSLPLGVHNSTLQ
VSDVLDKVRDLKSLSTNQLRSLVGLNLEGNGVATDVPSATKRWGFRSPVPPVNVYEG
EWAECYNLEIKKPDGSECLPAAPDGIKRGPRYVHKVSGTGCAGDFAFHKGAF
LYDLASTVLYRGTTFAEGVVAFILPQAKKDPFSSHPLREPVNATEDPSSGYSTTI
RYQATGTEVETELFEVDNLTVQLSSEBTPQLLOLNETIYTSKRSNTTGLIKW
VPEIDTIGWAFWETKCNLRKIRSELSFTVVSNGAKNISQSPARTSSDPGNT
TTEDHKINAGNSAMVQVHSQGREAAVSHLTITATISTSPQSTTKPGPNSTHTP
VYKLDISEATQVEQHRRTDNDSTPSATTAAGPPKAEKNTTSTKSTDFLDPATTT
SPONSHETAGNNTHODTGEBSASGKGLITNTIAGVAGLITGRRTRREALVNAQ
PKCNPLNHYMTTQDEGAIGLAWIPYFPGAEGVYIEGLMNDGLICGLRLANETT
QALQLFRATETLPTDFSLNRKAIDFLLQWRMGTCCHILGPCCCTEPHDWTNKTIDK
QIILHDFVDKTLPTDQGDNDNMWGTWRQMPAGIVGTGVIIVIALFCICKFVF"
7529..7540
/gene="GP"
/feature="encodes the glycoprotein cleavage site, precursor
GP is cleaved by subtilisin-like cellular protease furin
into subunits GP1 and GP2 that are linked by a disulfide
bond"
/citation=[3]

```

```

Query Match      66.1%; Score 1489.4; DB 14; Length 18959;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 ATGGCGTTACAGGAATATTGCAAGTTACCTCGTATCGAATCAAGAGGACATCATTTCTTT 60
Db |||||
6039 ATGGCGTTACAGGAATATTGCAAGTTACCTCGTATCGAATCAAGAGGACATCATTTCTTT 6098
QY 61 CTTTGGGTAAATTATCCTTTTCCAAGAAACATTTTCATCCCATCTTGGAGTCAATCACAAT 120
Db |||||
6099 CTTTGGGTAAATTATCCTTTTCCAAGAAACATTTTCATCCCATCTTGGAGTCAATCACAAT 6158
QY 121 AGCACATTACAGGTTAGTAGTATGTCGACAAATAGTTTGTGTCGACAAATCTGTCACCA 180
Db |||||
6159 AGCACATTACAGGTTAGTAGTATGTCGACAAATAGTTTGTGTCGACAAATCTGTCACCA 6218
QY 181 AATCAATTAGATAGTTGGAATCTCGAAGGGAATGGAATGGAATGGAATGGAATGGAATGGA 300
Db |||||
6219 AATCAATTAGATAGTTGGAATCTCGAAGGGAATGGAATGGAATGGAATGGAATGGAATGGA 6278
QY 241 TCTGCAACTAAAGATGCGGCTTCAAGTCCGCTGTCGACCAAAAGGTGCTCAATTATGAA 300
Db |||||
6279 TCTGCAACTAAAGATGCGGCTTCAAGTCCGCTGTCGACCAAAAGGTGCTCAATTATGAA 6338
QY 301 GCTGGTGAATGGCTGAAACTGCTACAATCTTGAATCAAAATACTGACCGGAGTGAG 360
Db |||||
6339 GCTGGTGAATGGCTGAAACTGCTACAATCTTGAATCAAAATACTGACCGGAGTGAG 6398
QY 361 TGTCTACCAAGCAGCGCAGACGGGATTCGGGGCTTCCCGGTCGGGTATGTCACAAA 420
Db |||||
6399 TGTCTACCAAGCAGCGCAGACGGGATTCGGGGCTTCCCGGTCGGGTATGTCACAAA 6458
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGGCTTCCATAAAGAGGTCCTTCTTC 480
Db |||||
6459 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGGCTTCCATAAAGAGGTCCTTCTTC 6518
QY 481 CTGTATGATCGACTGCTTCCACAGTTATCTACCGAGGAACGACTTTGCTCAAGGTGTC 540
Db |||||
6519 CTGTATGATCGACTGCTTCCACAGTTATCTACCGAGGAACGACTTTGCTCAAGGTGTC 6578
QY 541 GTTGCAATTTCTGATCTGCTGCAAGCTAAGAGGACTTCTCAGCTCAACACCTTGAGA 600
Db |||||
6579 GTTGCAATTTCTGATCTGCTGCAAGCTAAGAGGACTTCTCAGCTCAACACCTTGAGA 6638
QY 601 GAGCGGTCATGCAACGGAGGACCGTCTAGTGGCTACTATTCTACCAATATGATAT 660
Db |||||
6639 GAGCGGTCATGCAACGGAGGACCGTCTAGTGGCTACTATTCTACCAATATGATAT 6698
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTGAGGTTTCACAAATTTGACC 720
Db |||||
6699 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTGAGGTTTCACAAATTTGACC 6758
QY 721 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db |||||
6759 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 6818
QY 781 TATACAGTGGGAAAGGAGCAATACACCGGAAACCTAATTTGGAGGTCAACCCCGAA 840
Db |||||
6819 TATACAGTGGGAAAGGAGCAATACACCGGAAACCTAATTTGGAGGTCAACCCCGAA 6878
QY 841 ATTGTACAACTCGGGAGTGGGCTTCTGGAACCTAATAAATACTCACTAGAAAA 900
Db |||||
6879 ATTGTACAACTCGGGAGTGGGCTTCTGGAACCTAATAAATACTCACTAGAAAA 6937
QY 901 ATTCCAGTGAAGAGTTGTCTTTTCAAGTTGATCAACCGGAGCCAAAACATCAGTGT 960
Db |||||
6938 ATTCCAGTGAAGAGTTGTCTTTTCAAGTTGATCAACCGGAGCCAAAACATCAGTGT 6997
QY 961 CAGAGTCCGGCGGAACCTTCTCCGACCCAGGACCAACAACTCTGAAGACCAAAA 1020
Db |||||
6998 CAGAGTCCGGCGGAACCTTCTCCGACCCAGGACCAACAACTCTGAAGACCAAAA 7057
QY 1021 ATCATGGCTTCAGAAATCTCTGCAATGTTTCAAGTGCACAGTCAAGAGGAGGAGCT 1080
Db |||||
7058 ATCATGGCTTCAGAAATCTCTGCAATGTTTCAAGTGCACAGTCAAGAGGAGGAGCT 7117
QY 1081 GCAGTGTGCACTCTAACAAACCTTGGCAACAATCTCCACGAGTCCCAATCCCTCAACACC 1140

Db |||||
7118 GCAGTGTGCACTCTAAACAACCTTGGCCCAATCTCCAGAGTCCCAATCCCTCAACACC 7177
QY 1141 AAACAGAGTCCGGCAACAGCAGCAGCCCAATATACACCGGTGTATAACTTTGACATCTCTGAG 1200
Db |||||
7178 AAACAGAGTCCGGCAACAGCAGCAGCCCAATATACACCGGTGTATAACTTTGACATCTCTGAG 7237
QY 1201 GCAACTCAAGTTGAACAAACATCACCAGCAGAGCAGAAACAGCAGCAGCAGCAGCAGCAGCAGC 1260
Db |||||
7238 GCAACTCAAGTTGAACAAACATCACCAGCAGAGCAGAAACAGCAGCAGCAGCAGCAGCAGCAGC 7297
QY 1261 CCCTCTGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1320
Db |||||
7298 CCCTCTGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 7357
QY 1321 ACTGACTTCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1380
Db |||||
7358 ACTGACTTCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 7417
QY 1381 AACAAACACACTCATCACCAGAGATACCGAGAGAGAGTGCCAGCAGCGGGAGGCTAGGC 1440
Db |||||
7418 AACAAACACACTCATCACCAGAGATACCGAGAGAGAGTGCCAGCAGCGGGAGGCTAGGC 7477
QY 1441 TTAATTACCAATATCTATTGCTGAGTCCGAGAGCTGATCAACAGCGGGAGAGAACTCGA 1500
Db |||||
7478 TTAATTACCAATATCTATTGCTGAGTCCGAGAGCTGATCAACAGCGGGAGAGAACTCGA 7537
QY 1501 CGA 1503
Db |||||
7538 AGA 7540

RESULT 9
AF272001
LOCUS Zaire Ebola virus strain Mayinga, complete genome. VRL 28-AUG-2002
DEFINITION AF272001
ACCESSION AF272001.1 GI:11761745
VERSION
KEYWORDS
SOURCE Zaire ebolavirus (ZEBOV)
ORGANISM Zaire ebolavirus
Viruses; serona negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE 1 (bases 1 to 18959)
AUTHORS Volchkov, V.E., Volchkova, V.A., Chepurinov, A.A., Blinov, V.M.,
Dolnik, O., Netesov, S.V. and Feldmann, H.
TITLE Characterization of the L gene and 5' trailer region of Ebola virus
J. Gen. Virol. 80 (Pt 2), 355-362 (1999)
JOURNAL
MEDLINE 99171703
PUBMED 10073695
REFERENCE 2 (bases 1 to 18959)
AUTHORS Volchkov, V.E., Chepurinov, A.A., Volchkova, V.A., Ternovoj, V.A. and
Klenk, H.D.
TITLE Molecular characterization of guinea pig-adapted variants of Ebola virus
JOURNAL Virology 277 (1), 147-155 (2000)
MEDLINE 20517553
PUBMED 11062045
REFERENCE 3 (bases 1 to 18959)
AUTHORS Volchkov, V.E.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2000) Institute of Virology, Robert-Koch-Str.17,
Marburg 35037, Germany
FEATURES
source location/Qualifiers
1..18959 /organism="Zaire ebolavirus"
/viral /viral
/mol_type="genomic RNA"
/strain="Mayinga"
/db_xref="taxon:186538"
/note="subtype: Zaire; biotype: Guine pig adapted variant
8mc"
misc_feature 1..55


```

/feature="putative leader region"
56..3026
/gene="NP"
misc_feature
56..3026
/gene="NP"
56..67
/gene="NP"
/feature="transcription start signal"
470..2689
/gene="NP"
/feature="transcription start signal"
470..2689
/codon_start=1
/product="nucleoprotein NP"
/protein_id="AAG40164.1"
/db_xref="GI:11761746"
/translation="MDSRPQKIMWAPLSTESDMYHKILTAGLSVQQGIIVRQVPIVY
QVNNLEETQIIQAFEGVDFQSADSLMLCLHAYQDGYKLFLESGAVKYLEGH
GPRVEKRDGKLELLPAVSGGKNIKRTLAAMPBEEETEANAGOLFSLFSLFKP
LVGEKACLEKVRQIOVHAREGLIOVPTAQSGVHMVIPLMARTNFIKFLIHOG
MHVAGHDANDAVISNSVAQARFSGLLIVKTVDLHILQXTERGVRLHPLARTAKVNE
VNSFAULSLAKHGEYAPFARLLNLGNVLEHGLFPQLSAIALGAVATAGSTLAV
NVGEYQQLREAAEAEKQLOQYAESRELDHLGDQEKKILMNFHOKNBSIFQQTN
AMVTLRKERLAKLEATAASLPKTSQGHDDDDIPFGPINDDDNFGHQDDPDSQ
DTTIPDVVDDPQDSGYQSYQSENGMNPDDLVLFDLDEDDTKPVNRSKGGQ
KNSQKGOHIEGRQIQPIQNVPGPHRTIHASAPLTNDNRNPSGSTSPMLTPTN
BEAPLDADDETSPLPESDDEQDDGTSNRTPTVAPAPVYRDHSEKELPQDE
CQDQHTQEARNQSDNTQSEHSLFEMTHILRSQGPDAVLVTHMMKDEPVPVSTSD
QKETYDPDSEEBEYPPMLTEKAMNEENRFTLDGQQPYWPMNHKNKFMALQHQH"
3015..3026
/gene="NP"
3027..3031
/feature="intergenic region"
3032..4407
/gene="VP35"
3032..4407
/gene="VP35"
3032..4043
/gene="VP35"
/feature="transcription start signal"
3129..4151
/gene="VP35"
/feature="second nucleoprotein"
/codon_start=1
/product="VP35"
/protein_id="AAG40165.1"
/db_xref="GI:11761747"
/translation="MTTRTKRGHTAATTONDRMPGPELSGWISEQLMTGRIPVSDIF
CDIENPGCLVASQMQTKPNKPTNSOTQDPICNHSEFVQOTLASLATVVOQOTI
ASELEQRTISLENGLEKPYDMKTISSLNRVCAEMVAKYDLLWTTGRATATAATE
AYWABHGQPPGSLYEESAIRKIESRDETVPQSVREAFNNLSITSLTEENFGKPD
ISAKDLRIMYDHLPGFGTAHQVQLVICKLQKDSNLDIHAEFQFSAEGSDPQCA
LIQITKRVPIFDQAAPPVHIRSGDIPRACQKSLRPVPPSPKIDRGWVCVFQLDQDK
TLGLKI"
4390..5894
/gene="VP40"
4390..5894
/gene="VP40"
4390..4401
/gene="VP35"
/feature="transcription start signal"
4397..4407
/gene="VP35"
4479..5459
/gene="VP40"
/feature="membrane-associated"
/codon_start=1
/product="VP40"
/protein_id="AAG40166.1"
/db_xref="GI:11761748"
/translation="MRRVILTPAPPEYMEAIYPVRSNSTIARGGNSNTGFLTPRSVNG
DTPSNPLRIADTTIDHRSVTSYSSAFLEAMVNIISGPKVMKQIPIWPLGVAD
QKTSYFSDSTAAIMLASVTIHPGKATNPLVRVRLGPGIDPHPLRLIRIGNQAFLOE
FVLPPQYFPDLPOLTKLITQELPAATWTDTPGTSGNALRPGISFHPKLPILL
PNKSGKGNASDLTSPEKIQAIMTSLQDFKIVPIDPTKNIHMGIEVPTLVHKLTKGKY
TSKNGQPIIPVLPKYIGLDPVAPGDLTWITQDCDCHSPASLPAVIEK"
5883..5894
/gene="VP40"
5895..5899
/feature="intergenic region"
5900..8305
/gene="Gp"
5900..8305
/gene="Gp"
5900..5911
/gene="Gp"
/feature="transcription start signal"
join(6039..6923,6923..8068)
/feature="Gp"
/feature="vtrion spike glycoprotein; ribosomal slippage"
/codon_start=1
/product="surface glycoprotein Gp"
/protein_id="AAG40168.1"
/db_xref="GI:11761750"
/translation="MGVTGILQIPDRPKRTSPFLWVILLFORTFSIPLGVHNSLTQ
VSDVDKLVCRKLSSTNQLRSVGLNLEGVATDVPATKRWGRSPGPPKVVNYEAG
EWAENCYNLEIKKPDGSECLPAAPDGIIRGPRCRYVHKVSGTGACDGFAPKESGAF
LYDLASTVIYRGTTFAEGVAVFLLPOAKKDPFSSHPLEPNVATDPSGGYSTTI
RYQATGVTNETEYLFVNDLTVVQLESRTTPQLLNLETIYTSGRKSNITGKLIWK
VNEIDTIGEMAFWETKNLTKIRSELSFTVNSGAKNISQSPARTSSDRGNT
TTEDHKISENSAMVQVHSQGREAAVSHLTLTILSTSPSLITKPGDNDSTNP
VYKLDISEATQVQHQHRTDNDSTSPATTAAGPKKANTNTSKTDFLDPATT
SPQHSSTAGNNTTHQDTGEASGKGLITNTIAGVAGLITGSRTRREALVNAQ
PKCNPLHYWTQDEGAAGIAMPYFGPAAGIYIEGLMHNQDGLICGLRLANETT
QALQLPLRATTELRTFSILNRKADIFLLQRWGGTCHILGPDCCIIEPHDWTNITDKID
QIHDVFVDKTLDPQGDNDNWTGWRQIPAGIVGTGVIIVIALFCICKVP"
6039..7133
/gene="Gp"
/feature="secreted, non-structural"
/codon_start=1
/product="small non-structural, secreted glycoprotein sGP"
/protein_id="AAG40167.1"
/db_xref="GI:11761749"
/translation="MGVTGILQIPDRPKRTSPFLWVILLFORTFSIPLGVHNSLTQ
VSDVDKLVCRKLSSTNQLRSVGLNLEGVATDVPATKRWGRSPGPPKVVNYEAG
EWAENCYNLEIKKPDGSECLPAAPDGIIRGPRCRYVHKVSGTGACDGFAPKESGAF
LYDLASTVIYRGTTFAEGVAVFLLPOAKKDPFSSHPLEPNVATDPSGGYSTTI
RYQATGVTNETEYLFVNDLTVVQLESRTTPQLLNLETIYTSGRKSNITGKLIWK
VNEIDTIGEMAFWETKNLTKIRSELSFTVNSGAKNISQSPARTSSDRGNT
QKTKTSKQKIPQLQWFKCTVKEGKLCRI"
6918..6924
/gene="Gp"
/feature="Gp gene editing site; additional A residues are
inserted or deleted during transcription of the GP gene by
the viral polymerase that resulted in a reading frame
shift"
6999..7010
/gene="Gp"
/feature="encodes the glycoprotein sGP cleavage site (RVRR),
precursor sGP (presGP) is cleaved by subtilisine-like
proteases into two distinct subunits sGP and short carboxy
terminal glycopeptide (Delta peptide)"
7529..7540
/gene="Gp"
/feature="encodes the glycoprotein GP cleavage site (RTRR),
precursor GP (presGP) is cleaved by subtilisine-like
proteases into subunits GP1 and GP2 that are linked by a
disulfide bond"
7793..7870
/gene="Gp"
/feature="putative immunosuppressive motif"
7988..9740
/gene="VP30"
7988..8053
/gene="Gp"
/feature="transmembrane anchor"
8288..9740
/gene="VP30"

```

```
misc_feature      8288..8299
/gene="VP30"
/feature="transcription start signal"
polya_signal      8295..8305
/gene="VP30"
CDS               8509..9375
/gene="VP30"
/note="nucleoprotein"
/codon_start=1
/product="VP30"
/protein_id="AAG40169.1"
/db_xref="GI:11761751"
/translation="MEASLYERPRRAQRHSDHHDHVRASSRRENYRGEYRQSR
ASQVRPTVFHKRVRPLTPAPKDIPTLKLKPLCDSSFCCKKHQLESITDRLLL
LIARNTCSVEQLNITAPKDSRLNPATDDPFQEEGPKITLLTIKAEHWARQDIR

Query Match      66.1%; Score 1489.4; DB 14; Length 18959;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATGGGCGTTACAGGAATATTGCGATTACCTCTGATCGATTCAAGAGGACATCTTTT 60
DB 6039 ATGGGCGTTACAGGAATATTGCGATTACCTCTGATCGATTCAAGAGGACATCTTTT 6098
QY 61 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCAAAT 120
DB 6099 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCAAAT 6158
QY 121 AGCACATTCAGGTTAGTGTGATGTCGACAACTAGTTTGTGTCGACAACTGTCATCCACA 180
DB 6159 AGCACATTCAGGTTAGTGTGATGTCGACAACTAGTTTGTGTCGACAACTGTCATCCACA 6218
QY 181 AATCAATTCAGATCAGTTCGATCGAATCTCGAAGGAATGGAGTGGCACTGACGTGCCA 240
DB 6219 AATCAATTCAGATCAGTTCGATCGAATCTCGAAGGAATGGAGTGGCACTGACGTGCCA 6278
QY 241 TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCCAACCAAGGTGGTCAATATGAA 300
DB 6279 TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCCAACCAAGGTGGTCAATATGAA 6338
QY 301 GCTGTGAATGGCTGAAAACCTGCTCAATCTTGAATCAAAAACCTGACGGGAGTGAG 360
DB 6339 GCTGTGAATGGCTGAAAACCTGCTCAATCTTGAATCAAAAACCTGACGGGAGTGAG 6398
QY 361 TGTCTACAGACGCGCAGACGGGATTCGGGGCTTCCCGGTCGCGGTATGTGCACAAA 420
DB 6399 TGTCTACAGACGCGCAGACGGGATTCGGGGCTTCCCGGTCGCGGTATGTGCACAAA 6458
QY 421 GTATCAGGAACGGGACCGGTGTGCGGAGACTTTTGCCTTCCATAAAGAGGGTGTCTTCTTC 480
DB 6459 GTATCAGGAACGGGACCGGTGTGCGGAGACTTTTGCCTTCCATAAAGAGGGTGTCTTCTTC 6518
QY 481 CTGTATGATCGACTTCTTCCAGTTATCTACCGAGGAACGACTTTCCTGAAGGTGTC 540
DB 6519 CTGTATGATCGACTTCTTCCAGTTATCTACCGAGGAACGACTTTCCTGAAGGTGTC 6578
QY 541 GTTGCATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCAAGCTTCACACCCCTTGAGA 600
DB 6579 GTTGCATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCAAGCTTCACACCCCTTGAGA 6638
QY 601 GAGCCGGTCAATGCAACGGAGGACCGTCTAGTGGCTACTATTCTACCAATATAGATAT 660
DB 6639 GAGCCGGTCAATGCAACGGAGGACCGTCTAGTGGCTACTATTCTACCAATATAGATAT 6698
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTGAGGTTGACAAATTTGACC 720
DB 6699 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTGAGGTTGACAAATTTGACC 6758
QY 721 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
DB 6759 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 6818
QY 781 TATACAAGTGGGAAAAGGAGCAATACCCAGGGAAAACTAATTTGGGAAGGTCAACCCCGAA 840
```

```
DB 6819 TATACAAGTGGGAAAAGGAGCAATACCAACCGGAAAACCTAATTTTGGAAAGGTCAACCCCGAA 6878
QY 841 ATTGATACAAACATCGGGAGTGGGCTTCTTGGGAAACTTAAATAAAAACTCTACTAGAAAA 900
DB 6879 ATTGATACAAACATCGGGAGTGGGCTTCTTGGGAACT-AAAAAACTCTACTAGAAAA 6937
QY 901 ATTCGAGTGAAGAGTTTCTTTTCAAGTTGTATCAAAACGGAGCCAAAAACATCAGTGT 960
DB 6938 ATTCGAGTGAAGAGTTTCTTTTCAAGTTGTATCAAAACGGAGCCAAAAACATCAGTGT 6997
QY 961 CAGAGTCCGGGCGGAACTTCTTCCGACCCAGGAGCAACAACTGAGAGACACAAA 1020
DB 6998 CAGAGTCCGGGCGGAACTTCTTCCGACCCAGGAGCAACAACTGAGAGACACAAA 7057
QY 1021 ATCATGGCTTCAGAAAAATCTCTGCAATGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
DB 7058 ATCATGGCTTCAGAAAAATCTCTGCAATGTTCAAGTGCACAGTCAAGGAAGGAAGCT 7117
QY 1081 GCAGTGTGCGCATCTAAACAACTTTCGCCAATCTCCAGAGTCCCAATCCCTCAACCC 1140
DB 7118 GCAGTGTGCGCATCTAAACAACTTTCGCCAATCTCCAGAGTCCCAATCCCTCAACCC 7177
QY 1141 AAACAGGTCCGGACAAACAGACCCATTAATACACCCGTGTATAAATTGACATCTCTGAG 1200
DB 7178 AAACAGGTCCGGACAAACAGACCCATTAATACACCCGTGTATAAATTGACATCTCTGAG 7237
QY 1201 GCACCTAAGTTGAACAACTACCGCAGAACAGACAAACAGACGACGCTCCGACACT 1260
DB 7238 GCACCTAAGTTGAACAACTACCGCAGAACAGACAAACAGACGACGCTCCGACACT 7297
QY 1261 CCCTCTGCGACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACAGAGCAAGAGC 1320
DB 7298 CCCTCTGCGACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACAGAGCAAGAGC 7357
QY 1321 ACTGACTTCTTGACCCCGCCACCAACAAAGTCCCAAAACCAACAGAGGAGACCGCTGSC 1380
DB 7358 ACTGACTTCTTGACCCCGCCACCAACAAAGTCCCAAAACCAACAGAGGAGACCGCTGSC 7417
QY 1381 AACACAACTCTATCACCAGATACCGGAGAGAGTCCAGCAGCGGGAAGCTTAGGC 1440
DB 7418 AACACAACTCTATCACCAGATACCGGAGAGAGTCCAGCAGCGGGAAGCTTAGGC 7477
QY 1441 TTAATTACCAATCTATTGCTGGAGTCCAGGACTGATCAGAGCGGGGAGAAAGACTCGA 1500
DB 7478 TTAATTACCAATCTATTGCTGGAGTCCAGGACTGATCAGAGCGGGGAGAAAGACTCGA 7537
QY 1501 CGA 1503
DB 7538 AGA 7540

RESULT 10
AXI42960
LOCUS      18959 bp      RNA      linear      VRL 16-OCT-2002
DEFINITION Zaire Ebola virus strain Mayinga subtype Zaire, complete genome.
ACCESSION  AY142960
VERSION     AY142960.1  GI:23630482
KEYWORDS
SOURCE      Zaire ebolavirus (ZBOV)
ORGANISM    Zaire ebolavirus
            Viruses; ssRNA negative-strand viruses; Mononegavirales;
            Filoviridae; Ebola-like viruses.
REFERENCE   1 (bases 1 to 18959)
AUTHORS     Wilson,J.A., Kondig,J.P., Kuehne,A.I. and Hart,M.K.
TITLE       Ebola virus subtype Zaire strain Mayinga complete genome
JOURNAL     Unpublished
AUTHORS     Wilson,J.A., Kondig,J.P., Kuehne,A.I. and Hart,M.K.
TITLE       Direct Submission
JOURNAL     Submitted (16-AUG-2002) Virology, USAMRIID, 1425 Porter St.,
            Frederick, MD 21702, USA
FEATURES
            Location/Qualifiers
```


AADIFSRPSGKRRLQILGYLEGRTRLLASKIINNNTETPVLDRLKRLTLQRLWSLWFSY
LDQCNILAEALTOITCTVDLAQILREYSWAHLERPLIGATLPCMEQKFWLKP
YQCQCNSAKQPKPFSVAVKKHIVSMAWNASRISWTIGDGIPIYGSRTEDKIGQ
PAIKPCPSAARLEARELILASRLTWTWQSSNSDLILKPLEARVNLVSQEILOMTPSH
YSGNIVHRNDQYS PHS PANRMSNATRLIVSTWTLGFEFGGQSADSNLIIFQNV
NTAVALFDIKFRNTSATDIOYNRAHLHLTKCCTREVPAQILYITLTDLDLKYRENE
LIYDNPLKGGNLCNISENPFQKRLNIIEDDLIRPLHSGWELAKTIMQSIIDS
NNSSDPISSSTREFTTHFLYPIKILGLYSFAYSVYLVNTILTKLTLNDNPLY
L7TQIHNIPHRSLIKLPFKHASVMSRLMSIDPHFSYVIGGAAGDGLSDAARLPLR
TSISFLPTKVEWILNRTIVPLWIVYPLEGONPTPVNNFLVQIVELAVHSSRQAF
KTTISDHVPHDNLVYTCCKSTASNPFLASLAWRSRHRNSRKLARSDSTGSSNNNS
DGHIERQSDTRPDHGTENLVLQHEIKRITIPQENTHQGSPFOSFLSDSACGT
ANPKLFORSNRVDFQDHSNAREHQIISHRLVLPFFLTSQOTROUTSNESQTO
DEISIKYOLRLVDITVYFRTGIVSSMHYKLDVLEWIEBSFKSANDVLAEGEGAL
LLIKYQVKTFLFNTLATVSEISVIGMTVMMLPMSKFNQDLEIILNNSASQ
L7DITNTPWFKDQRLPKQVEVITMDAETENINRSKLYEAVYKLIHLIDPSVLKA
VVLKFLSDTEGMHLNDNLAPFPATGYLTKPITSSARSSEWYLCITFLSTTRKMPH
QNHLSCKVITLALQIQORSYWLSHLTYADCSHLUSYIRLGFPSLEKYLHRYNL
VDSKGPLVSIITQHLAHLRAEIRELTNDYNQORSTQTYHPRTAKGRITKLVDNDYL
KFLVLQALKNGTWOAEFKLPPELISVCNRFYHIDNCBERFLVQTLVYLHRMQDSE
VKLIERLTGLLSLFPDGLYRFD"

ORIGIN

Query Match 66.1%; Score 1489.4; DB 14; Length 18959;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 ATGGGGTTTACAGGAATATTGCAGTTTACCTCGTGTGATTCAGTTCAGAGGACATCTCTTT 60
DB 6039 ATGGGGTTTACAGGAATATTGCAGTTTACCTCGTGTGATTCAGTTCAGAGGACATCTCTTT 6098
QY 61 CTTTGGGTAAATATCTCTTTTCCAAAGAAATTTTCCATCCCACTGGAGTCATCCACAAT 120
DB 6099 CTTTGGGTAAATATCTCTTTTCCAAAGAAATTTTCCATCCCACTGGAGTCATCCACAAT 6158
QY 121 AGCATTATACAGTTAGTGTGATGTCGACAAAATAGTTTGTGTCGACAAAATCTGATCCACA 180
DB 6159 AGCATTATACAGTTAGTGTGATGTCGACAAAATAGTTTGTGTCGACAAAATCTGATCCACA 6218
QY 181 AATCAATTACAGTTCAGTTCGATGTAATCTCGAAGGGAATGAGTGCACACTGACGTGCA 240
DB 6219 AATCAATTACAGTTCAGTTCGATGTAATCTCGAAGGGAATGAGTGCACACTGACGTGCA 6278
QY 241 TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCCAACCAAGGTGCTCAATATATGAA 300
DB 6279 TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCCAACCAAGGTGCTCAATATATGAA 6338
QY 301 GCTGTGTAATGGGCTGAAAATCTGCTACAATCTTGAATCAAAAACCTGACGGGAGTGAG 360
DB 6339 GCTGTGTAATGGGCTGAAAATCTGCTACAATCTTGAATCAAAAACCTGACGGGAGTGAG 6398
QY 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTTCGGGTATCTGTCACAAA 420
DB 6399 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTTCGGGTATCTGTCACAAA 6458
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCATATAAGAGGGTGTCTTCTTC 480
DB 6459 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCATATAAGAGGGTGTCTTCTTC 6518
QY 481 CTGTATGATCGACTTGTCTTCCAGTATCTACCGAGGAACGACTTTGCTGAAAGGTGTC 540
DB 6519 CTGTATGATCGACTTGTCTTCCAGTATCTACCGAGGAACGACTTTGCTGAAAGGTGTC 6578
QY 541 GTTGCATTTCTGATCTGCTCCCAAGCTAAGAGGACTTCTTCACTGCTCACACCCCTTGAGA 600
DB 6579 GTTGCATTTCTGATCTGCTCCCAAGCTAAGAGGACTTCTTCACTGCTCACACCCCTTGAGA 6638
QY 601 GAGCCGGTCAATGCAACGGGAGACCGGTCTAGTGGCTACTATTCTACCAACAATTAGATAT 660
DB 6639 GAGCCGGTCAATGCAACGGGAGACCGGTCTAGTGGCTACTATTCTACCAACAATTAGATAT 6698
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCAGGTTGACAAATTTGACC 720

DB 6699 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTGACAAATTTGACC 6758
QY 721 TACGTCCAACTTGAATCAAGATTTACACACACAGTTCCTGCTCCAGCTGAATGAGACAATA 780
DB 6759 TACGTCCAACTTGAATCAAGATTTACACACACAGTTCCTGCTCCAGCTGAATGAGACAATA 6818
QY 781 TATACAAGTGGGAAAGAGGACAAATACCAACGGGAAATCTAATTTGGAAAGTCAACCCCGAA 840
DB 6819 TATACAAGTGGGAAAGAGGACAAATACCAACGGGAAATCTAATTTGGAAAGTCAACCCCGAA 6878
QY 841 ATTGATACAAACCAATCGGGAGTGGGCTTCTGGGAACTATAAAAAAATCTACTAGAAAA 900
DB 6879 ATTGATACAAACCAATCGGGAGTGGGCTTCTGGGAACTATAAAAAAATCTACTAGAAAA 6937
QY 901 ATTCGAGTGAAGAGTTCCTTTTACAGTTGTATCAAAACGGGACCAAAAAACATCAGTGTG 960
DB 6938 ATTCGAGTGAAGAGTTCCTTTTACAGTTGTATCAAAACGGGACCAAAAAACATCAGTGTG 6997
QY 961 CAGAGTCCGGCGGAACTTTCTTCGACCCAGGGACCAACAAACAACTGAAGACACAAAA 1020
DB 6998 CAGAGTCCGGCGGAACTTTCTTCGACCCAGGGACCAACAAACAACTGAAGACACAAAA 7057
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1080
DB 7058 ATCATGGCTTCAGAAAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAGGGAAGCT 7117
QY 1081 GCAGTGTGCGCATCTAACAAACCTTGGCCAACTCTCCAGAGTCCCAATCCCTCACAACC 1140
DB 7118 GCAGTGTGCGCATCTAACAAACCTTGGCCAACTCTCCAGAGTCCCAATCCCTCACAACC 7177
QY 1141 AAACAGAGTCCGGACCAACAGCACCCCATTAATACACCGTGTATAAACTTGACATCTCTGAG 1200
DB 7178 AAACAGAGTCCGGACCAACAGCACCCCATTAATACACCGTGTATAAACTTGACATCTCTGAG 7237
QY 1201 GCACTCAAGTTGAACACATCATCCGAGAACAGACAGACAGACAGTCCGACACT 1260
DB 7238 GCACTCAAGTTGAACACATCATCCGAGAACAGACAGACAGACAGTCCGACACT 7297
QY 1261 CCCTCTGCGACAGCGGACCGGACCCCAAAAGCAGAGAACACCAACAGCAGCAGAGC 1320
DB 7298 CCCTCTGCGACAGCGGACCGGACCCCAAAAGCAGAGAACACCAACAGCAGCAGAGC 7357
QY 1321 ACTGATTTCTGGACCCCGCCACCAACAAAGTCCCCAAAAACCAAGCAGAGACCGTGGC 1380
DB 7358 ACTGATTTCTGGACCCCGCCACCAACAAAGTCCCCAAAAACCAAGCAGAGACCGTGGC 7417
QY 1381 AACACACACTCATCACAGATACCGAGAACAGAGTCCGACAGCGGGAGGTAGG 1440
DB 7418 AACACACACTCATCACAGATACCGAGAACAGAGTCCGACAGCGGGAGGTAGG 7477
QY 1441 TTAATTACCAATATCTATTGCTGGAGTCCGAGGACTGATCAGAGCGGGAGAGAACTCGA 1500
DB 7478 TTAATTACCAATATCTATTGCTGGAGTCCGAGGACTGATCAGAGCGGGAGAGAACTCGA 7537
QY 1501 CGA 1503
DB 7538 AGA 7540

RESULT 11

EVU81161
LOCUS
DEFINITION Zaire Ebola virus virion spike glycoprotein and small secreted non-structural glycoprotein sgp (GP) gene, complete cds.
ACCESSION U81161
VERSION U81161.1 GI:1753170
KEYWORDS Zaire ebolavirus (ZEBOV)
ORGANISM Zaire ebolavirus
SOURCE Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Ebola-like viruses.
REFERENCE 1 (sices)
AUTHORS Volchkov, V.E., Blinov, V.M., Kotov, A., Chempurnov, A. and Netesov, S.

TITLE The full-length nucleotide sequence of the Ebola virus
JOURNAL Thesis (1993) Institut f. Virologie, Philipps-Universitaet
REFERENCE 2 (sites)
AUTHORS Volchkov, V., Volchkova, V., Eckel, C., Klenk, H. D., Bouloy, M.,
 LeGuennou, B. and Feldmann, H.
TITLE Emergence of subtype Zaire Ebola virus in Gabon
JOURNAL Virology 232 (1), 139-144 (1997)
MEDLINE 97329084
PUBMED 9185597
REFERENCE 3 (bases 1 to 2408)
AUTHORS Volchkov, V.E.
TITLE Direct Submision
JOURNAL Submitted (06-DEC-1996) Volchkov V. E., Philipps-Universitaet,
 Institut f. Virologie, Robert-Koch-Str.17, Marburg, Germany, 35037
FEATURES Location/Qualifiers
 source
 1. .2408
 /organism="Zaire ebolavirus"
 /mol_type="genomic RNA"
 /strain="Eckron-76"
 /db_xref="taxon:186538"
 3. .13
 /function="transcriptional start signal for GP mRNA"
misc_signal
 142. .2171
 /gene="Gp"
 142. .2171
 /gene="Gp"
 /function="structural protein"
 /function="cell receptor recognition"
 /function="fusion"
 /notes="full-length glycoprotein synthesized from edited
 mRNA species; an additional A residue is inserted at the
 editing site during GP gene transcription"
 /codon_start=1
 /evidence=experimental
 /exception="RNA editing"
 /product="viral spike glycoprotein"
 /protein_id="AAC57993.1"
 /db_xref="GI:1753171"
 /translations="MGVTGILQLPRDRFKRTSFFLWVILFQRTFSIPILGVHNSTLQ
 VNDVDKLVCRDLSLQNLRSVGLNLEGVATDVPSATKRMGFRGVPKPVNYBAG
 EWAENCVNLEIKKPDGSECLPAAPDGI RGFPRCRYVHKVSGTGPCAGDPAFKEGAFP
 LYDRLASTVIYRGTTAEGVAVFLIILPQAKKDFSSHPLREPVNATEDSSGGYYSTTI
 RYQATGFTNETSLYFEVDNLTLYQLESRETPOFLIQLNETIYTSGRKSNNTTGLKIWK
 VNPEIDTTIGEMAFWETKTSLSKFAKSCLSQLYTEPKTSVVRVRELLPTQGPPTO
 TTEDHKIMASENSAMVQSGREAAVSHLTATISTSPQSLTTKPGPDNSTHTPT
 TTDKLI SEATQVSOHRRRTDNDTSADTSPSATTAAAGPKAENNTSKSDFLDPAATT
 SPQNSETAGNNNTHQDTEESASSKGLGLINTIAGVAGLITGRRTRREALVNAQ
 KQCNPLNHWITODEGATLAWIPYFGPAAGEIYTEGLMHNQNLGILGURQLANST
 PAKQLRLHATTEURLTSLNKLAWIDFLLQWGTGCTHILGPDCCIEBPHDWTNTDKID
 QIITHDFVDTLPDQGDNDNNWTTGWRQWIPAGIGVTGVIIAIVIALFCICKFVF"
 142. .1236
 /gene="Gp"
 /notes="synthesized from non-edited GP mRNA species."
 /codon_start=1
 /evidence=experimental
 /product="small secreted non-structural glycoprotein sgp"
 /protein_id="AAC57993.1"
 /db_xref="GI:1753172"
 /translations="MGVTGILQLPRDRFKRTSFFLWVILFQRTFSIPILGVHNSTLQ
 VNDVDKLVCRDLSLQNLRSVGLNLEGVATDVPSATKRMGFRGVPKPVNYBAG
 EWAENCVNLEIKKPDGSECLPAAPDGI RGFPRCRYVHKVSGTGPCAGDPAFKEGAFP
 LYDRLASTVIYRGTTAEGVAVFLIILPQAKKDFSSHPLREPVNATEDSSGGYYSTTI
 RYQATGFTNETSLYFEVDNLTLYQLESRETPOFLIQLNETIYTSGRKSNNTTGLKIWK
 VNPEIDTTIGEMAFWETKTSLSKFAKSCLSQLYTEPKTSVVRVRELLPTQGPPTO
 QLTKTSKWLQKILPQWPKCTVSGKLCQRI"
 2391. .2402
 /functions="transcriptionl start signal for VP30 mRNA" -
 2398. .2408
misc_signal
polya_signal
ORIGIN

```
QY 1081 GCAGTGTGCGCATCTAACCAACCCCTTGCCACAATCTCCACGAGTCCCAATCCCTCAACCC 1140
Db 1221 GCAGTGTGCGCATCTAACCAACCCCTTGCCACAATCTCCACGAGTCCCAATCCCTCAACCC 1280
QY 1141 AAACGAGTCCGGACACAGCACCCATATACACCGGTGTATAAACTTGACATCTCTGAG 1200
Db 1281 AAACGAGTCCGGACACAGCACCCATATACACCGGTGTATAAACTTGACATCTCTGAG 1340
QY 1201 GCAACTCAAGTTGAACAACATCACCGCAGAAACAGACAAACAGACAGACAGCCTCCGACACT 1260
Db 1341 GCAACTCAAGTTGAACAACATCACCGCAGAAACAGACAAACAGACAGACAGCCTCCGACACT 1400
QY 1261 CCCTCTGCCACAGCGCAGCGGACCGGACCCCAAGCAGAGAAACACCAACAGCAGCAGAGC 1320
Db 1401 CCCTCTGCCACAGCGCAGCGGACCGGACCCCAAGCAGAGAAACACCAACAGCAGCAGAGC 1460
QY 1321 ACTGACTTCTCGACCCCGCCACCAACAAAGTCCCAAAACACAGCGAGACCCCTGGC 1380
Db 1461 ACTGACTTCTCGACCCCGCCACCAACAAAGTCCCAAAACACAGCGAGACCCCTGGC 1520
QY 1381 AACAAACACTCATCACCAAGATACCGGAGAGAGAGTGCCAGCAGCGGGAAGCTAGGC 1440
Db 1521 AACAAACACTCATCACCAAGATACCGGAGAGAGAGTGCCAGCAGCGGGAAGCTAGGC 1580
QY 1441 TTAATTACCAATACTATTCTGCTGGAGTCGAGGACTGATCACAGCGGGAGAGAACTCGA 1500
Db 1581 TTAATTACCAATACTATTCTGCTGGAGTCGAGGACTGATCACAGCGGGAGAGAACTCGA 1640
QY 1501 CGA 1503
Db 1641 AGA 1643

RESULT 12
AF499101
LOCUS Zaire Ebola virus strain Mayinga, linear VRL 28-AUG-2002
DEFINITION Zaire Ebola virus strain Mayinga, complete genome.
ACCESSION AF499101
VERSION AF499101.1 GI:21702647
KEYWORDS Zaire ebolavirus (ZEBOV)
SOURCE Zaire ebolavirus
ORGANISM Viruses; sRNA negative-strand viruses; Mononegavirales;
REFERENCE 1 (bases 1 to 18960)
AUTHORS Wilson,J.A., Kondig,J.P., Kuehne,A.I. and Hart,M.K.
TITLE Mouse adapted variant of Ebola virus subtype Zaire strain Mayinga
complete genome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 18960)
AUTHORS Wilson,J.A., Kondig,J.P., Kuehne,A.I. and Hart,M.K.
TITLE Direct Submission
JOURNAL Submitted (05-APR-2002) Diagnostic System Division, USAMRIID, 1425
Porter St., Ft. Detrick, Frederick, MD 21702, USA
FEATURES
source
1..18960
/organism="Zaire ebolavirus"
/virion
/mol_type="genomic RNA"
/strain="Mayinga"
/db_xref="taxon:186538"
/notes="subtype: Zaire; biotype: mouse adapted variant"
misc_feature 1..55
/notes="putative leader region"
gene 56..3026
/gene="NP"
mRNA 56..3026
/gene="NP"
/misc_signal 56..67
/notes="nucleoprotein NP"
misc_signal 56..67
/notes="transcription start signal"
CDS 470..2689
```

```
/gene="NP"
/codon_start=1
/product="nucleoprotein NP"
/protein_id="AAM76031.1"
/db_xref="GI:21702648"
/translation="MDSRPQKIWMAPSLTESDMYHKILKTAGLSVQQGIIVRQRIIPVY
QVNLERICQLIIQAFAGVDVQBSADGFLMLCLHHAYQGDYKLFLESGAVKYLEGH
GFRFEVKRKGDLLELLPAVSSGKNIKRTLAAMPEEETTEANAGQFLSFASLFLPK
LVGEKACLEKVRQIOVHAQOGLIOVPTAQSVGHMMVIPLMRNTNFIKFLILHQG
MHVAGHADANDAVISNVAOARPSGLIIVKTLDHILLOKTERGVRLHPLARTAKVNE
VNSFKALSSLAKEGEYAPFARLLNLSGVNLEHGLFQLSALIALGVATAGSTIAGV
NVSGYOQLREATAEAEKQQYAESEHLDHGLDDQEKILMMFHQKNEISFQQTIN
AMVTLRKERLEAKTEATTAASLPKTSYGHYDDDDIPFPGINDDDPKVQDDDDTDSQ
DTTTPDVVVDGDSYGEYSYSGNNAPDOLVLFOLDDEDEDTKPVNRTSGGQQ
KNSQGOHIEGRQTSRPIQNVPGPHRTIHHASAPLTONDRNPSGTSRMLPTPIN
SEADPLDADDETSLLPLESDDEQDDEGTNRTPTPVAPPAVYRDHSEKKELPQDE
QQDDHTQEARNQSDNTQSEHSFEEMRYHILRSQGFDAVLIIYHMKDEPVVFTSD
GKETYTPDLSLEEYPPMLTEKAMNEENRFVTLGQQQFYPMVNMHNKFMALQHHQ"
/gene="NP"
3027..3031
/notes="intergenic region"
3032..4407
/gene="VP35"
3032..4407
/gene="VP35"
/product="VP35"
3032..3043
/gene="VP35"
/notes="transcription start signal"
3129..4151
/gene="VP35"
/codon_start=1
/product="VP35"
/protein_id="AAM76032.1"
/db_xref="GI:21702649"
/translation="MTTRTKRGHTVATTQNDRMGPGLSGWISQLMTGRIPVSDIF
CDIENNPGLCYASQMOQTKPNPKTRNSQTDPICNHSFEEVVTQSLATVVOQQTII
ASELSIORITSLNGLKPVYDMAKTISSLNRVCAEMVAKYDLLVMTTGRATATAATE
AYWAHGGPPGPPSLYEESAIRKIESRDETVPQSVREAFNNLSTTSITEENFGKPD
ISAKDLRNMIDHLPFGTGAHQVLQVQICKGKOSNSLDIIHAERQASLAEGDSQCA
LIQITKRVPIFQDAAPPVHIHRSKGDIPRACQSLRPPPPPSLDRGWVCVFLQDQGK
TLGLKI"
4390..5894
/gene="VP40"
4390..5894
/gene="VP40"
/product="membrane-associated matrix protein VP40"
4390..4401
/gene="VP40"
/notes="transcription start signal"
4397..4407
/gene="VP35"
4479..5459
/gene="VP40"
/codon_start=1
/product="membrane-associated matrix protein VP40"
/protein_id="AAM76033.1"
/db_xref="GI:21702650"
/translation="MRRVILPTAPPEYMEALPYVRSNSTIARGGNSNTGFLTPEVNG
DTPGNLPIADDTIDHASHTPGSVSSAFILEAMVNVISGPKVLMKQIPILWPLGVAD
QKTSYSDTTAAIMLASYTIHFQKATNPLVRVNRGLFGCIPDHPLELRILRIGNAFLQE
FVLVPQLPOVFTFDLTALKITQPLPAATWTDPTGSGALRPGISFHPKRLPLL
PNKSGKGSADLTSPEIKQAIMTSLQDPKIVIPIDPTKNIMGIEVPETLVKLTGKKV
TSKGCQPIIPVLLPKYIGLDPVAPGDLTMTVITQDCDCTCHSPASLPAVIEK"
5883..5894
/gene="VP40"
5895..5899
/notes="intergenic region"
5900..8305
/gene="GP"
5900..8305
/gene="GP"
/notes="transcription start signal"
```

misc_signal	/product="glycoprotein precursor" 5900..5911 /genes="GP" /notes="transcription start signal" join(6039..6923,6923..8068) /genes="GP" /notes="virion spike glycoprotein" /codon_start=1 /exception="RNA editing" /product="surface glycoprotein GP precursor" /protein_id="AAM76034.1" /db_xref="GI:21702651" /translation="MGVTGILQLPRDRFKTSFELWVILFORTFSLPLGVHNSTLQ VSDVDKLVCRKLSSTNQLRPVGLNLEGNVATDVPSATKRWGRFSGVPPKVNVEAG EWANCYNLEIKKPDGSECLPAADGIRGPRCRYVHKVSGTGPCAGDFAPFKEGAFF LYDLASTVIYRGTTFAGVVAFLILPOAKKDFSSHPLEPNVATDPSSGYSTTI RYOATGFTNTEYLFVDNLTYVQLPRFTPOFLQLNETIYTSGRKNTTGLIWK VNPEDITIGWAFWETKNTKRIRSELSFTVWNGAKNISQSPARTSSDPGTNT TTEDHKIMASINSAMQVHSQGREAAVSHLTTLATISTSPQSILTKPGDNTHTNP VYKLDISEATQVEHRRITDNDSTASDTPSATTAAGPPKAENTWNTSDFLDPATTT SPQNHSETANNTHHQDTGEESSAGSGKGLITNTIAGVAGLITGGRRTRREAIVNAQ PKCPNHYWTTQDEGAIGLAWIPYFGPAAEGIYTEGLMHNQDGLICGLRLANET QALQLFRATTELRTFSLNRKAIDFLQRWGGTCHILGPDCCIPEHDWTNINDKID QIIDFVDKTLDPDQSDNDNWNWGRQWIPAGIGVTGVIIAIVIALFCICKVF" 7529..7540 /genes="GP" /notes="site of cleavage of precursor GP, preGP, by subtilisin-like proteases into subunits Gpi and Gp2 that are linked by a bond; Region: glycoprotein GP precursor cleavage site" 7793..7870 /genes="GP" /notes="Region: putative immunosuppressive motif" 7988..8053 /genes="GP" /notes="Region: transmembrane anchor" 8039..7133 /genes="GP" /codon_start=1 /product="small non-structural, secreted glycoprotein sgp precursor" /protein_id="AAM76035.1" /db_xref="GI:21702652" /translation="MGVTGILQLPRDRFKTSFELWVILFORTFSLPLGVHNSTLQ VSDVDKLVCRKLSSTNQLRPVGLNLEGNVATDVPSATKRWGRFSGVPPKVNVEAG EWANCYNLEIKKPDGSECLPAADGIRGPRCRYVHKVSGTGPCAGDFAPFKEGAFF LYDLASTVIYRGTTFAGVVAFLILPOAKKDFSSHPLEPNVATDPSSGYSTTI RYOATGFTNTEYLFVDNLTYVQLPRFTPOFLQLNETIYTSGRKNTTGLIWK VNPEDITIGWAFWETKNTKRIRSELSFTVWNGAKNISQSPARTSSDPGTNT QKTKSWIQKIPQWFKCTVKEGKLQCR1" 6999..7010 /genes="GP" /notes="site of cleavage of sgp precursor by subtilisin-like proteases into two distinct subunits, sgp and short carboxy-terminal glycopeptide, delta peptide; Region: glycoprotein sgp precursor cleavage site" 6918..6924 /genes="GP"	polyA_signal 9730..9740 /gene="VP30" /gene="VP30" 9741..9884 /notes="intergenic region" 9885..11497 /gene="VP24"	Query Match 66.0%; Score 1486.2; DB 14; Length 18960; Best Local Similarity 99.7%; Pred. No. 0; Matches 1499; Conservative 0; Mismatches 3; Indels 1; Gaps 1;	Qy 1 ATGGCGCTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 60 Db 6039 ATGGCGCTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 6098
misc_feature			Qy 61 CTTTGGGTAAATTATCTTTTCCAAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAA 120 Db 6099 CTTTGGGTAAATTATCTTTTCCAAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAA 6158	
misc_feature			Qy 121 AGCACATTACAGGTTAGTGATGTCGACAACTAGTTTGTGTGTGTCGACAACTGTCATCCACA 180 Db 6159 AGCACATTACAGGTTAGTGATGTCGACAACTAGTTTGTGTGTGTCGACAACTGTCATCCACA 6218	
misc_feature			Qy 181 AATCAATTGAGATCAGTTGGACTGAAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 240 Db 6219 AATCAATTGAGATCAGTTGGACTGAAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 6278	
CDS			Qy 241 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCCCAACCAAGAGGTGGTCAATTATGAA 300 Db 6279 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCCCAACCAAGAGGTGGTCAATTATGAA 6338	
misc_feature			Qy 301 GCTGGTGAATGGGCTGAAAACTGTACAATCTTGAATCAAAAACTGACGGGAGTGAG 360 Db 6339 GCTGGTGAATGGGCTGAAAACTGTACAATCTTGAATCAAAAACTGACGGGAGTGAG 6398	
misc_feature			Qy 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGGTGCGGTATGTGCACAAA 420 Db 6399 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGGTGCGGTATGTGCACAAA 6458	
misc_feature			Qy 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 480 Db 6459 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 6518	
misc_feature			Qy 481 CTGTATGATCGACTTGTCTCCACAGTTATCTTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540 Db 6519 CTGTATGATCGACTTGTCTCCACAGTTATCTTACCGAGGAACGACTTTTCGCTGAAGGTGTC 6578	
gene			Qy 541 GTTGCAATTTCTGATCTAGTCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600 Db 6579 GTTGCAATTTCTGATCTAGTCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 6638	
mRNA			Qy 601 GAGCCGCTCAATGCAACGGAGGACCCGCTTAGTGGCTACTATTCTTACCACAATTAGATAT 660 Db 6639 GAGCCGCTCAATGCAACGGAGGACCCGCTTAGTGGCTACTATTCTTACCACAATTAGATAT 6698	
misc_signal			Qy 661 CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720 Db 6699 CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 6758	
polyA_signal			Qy 721 TACGTCCAACTTGAATCAAGATTCCACCAAGTTTCTGCTCCAGCTGAATGAGACATA 780 Db 6759 TACGTCCAACTTGAATCAAGATTCCACCAAGTTTCTGCTCCAGCTGAATGAGACATA 6818	

Qy	781	TATACAGTGGGAAAGAGCAATACACCGGAAACCTAATTTTGGAAAGTCAACCCGAA	840
Db	6819	TATACAGTGGGAAAGAGCAATACACCGGAAACCTAATTTTGGAAAGTCAACCCGAA	6878
Qy	841	ATTGATACAAACAATCGGGAGTGGCCCTCTGGGAAACTATAAAAAACCTCAGTACGAAA	900
Db	6879	ATTGATACAAACAATCGGGAGTGGCCCTCTGGGAAACT-AAAAAACTCAGTACGAAA	6937
Qy	901	ATTGCGAGTGAAGAGTTGTCTTTTACAGTTGTATCAAAACGAGGCCAAAAACATCAGTGGT	960
Db	6938	ATTGCGAGTGAAGAGTTGTCTTTTACAGTTGTATCAAAACGAGGCCAAAAACATCAGTGGT	6997
Qy	961	CAGAGTCCGGCGGAACTTCTTCGACCCAGGAGCAACAACAACCTGAAGACCAAAA	1020
Db	6998	CAGAGTCCGGCGGAACTTCTTCGACCCAGGAGCAACAACAACCTGAAGACCAAAA	7057
Qy	1021	ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCNAGTGCACAGTCAAGGAAGGAGCT	1080
Db	7058	ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCNAGTGCACAGTCAAGGAAGGAGCT	7111
Qy	1081	GCAGTGTGCATCTTAACAACCCCTTGGCCACAATCTCCACAGAGTCCCCAATCCCTCAACAAC	1140
Db	7118	GCAGTGTGCATCTTAACAACCCCTTGGCCACAATCTCCACAGAGTCCCCAATCCCTCAACAAC	7177
Qy	1141	AAACCGAGTCCGGACAAACAGACCCATATACACCCGTGTATAAACTTGACATCTCTGAG	1200
Db	7178	AAACCGAGTCCGGACAAACAGACCCATATACACCCGTGTATAAACTTGACATCTCTGAG	7237
Qy	1201	GCAACTCAAGTTGAAACAACATCACCAGACAGACAGACAGACAGACAGACAGACAGAC	1260
Db	7238	GCAACTCAAGTTGAAACAACATCACCAGACAGACAGACAGACAGACAGACAGACAGAC	7297
Qy	1261	CCCTCTGCCACGACCGCAGCGGACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC	1320
Db	7298	CCCTCTGCCACGACCGCAGCGGACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC	7357
Qy	1321	ACTGACTTCTGGACCCCGCCACACAAACAGTCCCAAAACCAACGAGACCGCTGGC	1380
Db	7358	ACTGACTTCTGGACCCCGCCACACAAACAGTCCCAAAACCAACGAGACCGCTGGC	7417
Qy	1381	AACAAACAACACTATACCAAGATACCGGAGAGAGAGTGCACAGCAGCGGAGAGTAGGC	1440
Db	7418	AACAAACAACACTATACCAAGATACCGGAGAGAGAGTGCACAGCAGCGGAGAGTAGGC	7477
Qy	1441	TTAATTACCAATACTATTGTCTGGAGTCCGAGGACTGATCAAGCGGGAAGAACTCGA	1500
Db	7478	TTAATTACCAATACTATTGTCTGGAGTCCGAGGACTGATCAAGCGGGAAGAACTCGA	7537
Qy	1501	CGA 1503	
Db	7538	AGA 7540	
RESULT 13			
AY354458		18961 bp ss-RNA linear	VRL 06-FEB-2004
LOCUS	AY354458		
DEFINITION	Zaire ebolavirus strain Zaire 1995, complete genome.		
ACCESSION	AY354458		
VERSION	AY354458.1	GI:33860540	
KEYWORDS			
SOURCE			
ORGANISM			
	Zaire ebolavirus (ZEBOV)		
	Zaire ebolavirus		
	Viruses; ssRNA negative-strand viruses; Mononegavirales;		
	Filoviridae; Ebola-like viruses.		
REFERENCE			
AUTHORS	1 (bases 1 to 18961)		
	Chain,P.S.G., Ichou,M.A., Malfatti,S.A., Hajjaj,A., Vergez,L.M.,		
	Paragas,J., Do,L.H., Jahrling,P.B., Smith,K.L., McCready,P.M. and		
	Ibrahim,M.S.		
TITLE	Zaire Ebola virus, complete genome		
JOURNAL	Unpublished		
REFERENCE			
AUTHORS	2 (bases 1 to 18961)		
	Chain,P.S.G., Malfatti,S.A., Hajjaj,A., Vergez,L.M.,		
	Smith,K.L. and McCready,P.M.		

TITLE	Direct Submission	
JOURNAL	Submitted (24-JUL-2003) Viral Sequencing Group, Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA	
REFERENCE	3 (bases 1 to 18961)	
AUTHORS	Ichou,M.A., Paragas,J., Jahrling,P.B., Ibrahim,M.S., Lofts,L., Hevey,M. and Schmaljohn,A.	
TITLE	Direct Submission	
JOURNAL	Submitted (24-JUL-2003) Viral Genomics Group, USAMRIID, Fort Detrick, MD 21702, USA	
FEATURES	Location/Qualifiers	
source	1..18961 /organism="Zaire ebolavirus" /mol_type="genomic RNA" /strain="Zaire 1995" /db_xref="taxon:186538"	
5'UTR	1..55 /note="putative leader region" /function="regulation or initiation of RNA replication"	
gene	56..3026 /gene="NP"	
mRNA	56..3026 /gene="NP" /product="nucleoprotein"	
misc_signal	56..67 /gene="NP"	
CDS	470..2689 /note="putative; transcription start signal" /gene="NP" /function="encapsidation of genomic RNA" /codon_start=1 /product="nucleoprotein" /protein_id="AAQ55045.1" /db_xref="GI:33860541"	
	translation="MDSRPQVMPTSLTESDMVDYHKILTAGLSVQOQIVRQRPVIVYQVNLEELCOLIQAEAGVDQESADSLLMLCLHAYQGDYKLFESGAVKLEHGGRFPEVKRDKGVQRLLEELPAVSSGKNIKRTLAMPEETTEANACQFLSPASLFLPKLVVGEKACUEQVQKQVQLHAPQSVGHMMVIFRLMRTNFLIKFLLIHOGMMVAAGLANDAVINSVAQAFSGLLIVKTVLDHTLOKTERGVRLHPRLARTAKVNEVSFAKSLAKHGAYAPFARLLNSGVNLEHGLFPQLSALGAVATAHSGTLAGVNGVEOQQLREATAEKOQVAESRELDHLGDQEKILANFHQKNEISFOQTNMYTLKERLAKLTENITAAISLPKTSGHVDDDDDIFFPGPINDDDHPGHDPTDSDTTIPVVPVDDGSGVGEYSKSENGMNAPDULVFLDDEDDETKPVNRSRTKGGQQKNSQKQGTGEGRTQSGTQSPQSDVPGPHRTIHASAPLTDNRNPESGTSPRMLTPIKEADPLDDADDTSSLPPELSDDEQDRGTNRTTAPVAFVDRHSKRLPQDEQDQHTQBARQDSQNTQPHSFSEMYRHILRSQGFDAVLVYHMKDEPVPVFSTSDGKEYTVPDSLESEYPPWLTEKEAMNEENRFVTLDDGGQFYWPVNVNHNKFWAILQHQQ"	
polya_signal	3015..3026 /gene="NP"	
gene	3032..4407 /gene="VP35"	
mRNA	3032..4407 /gene="VP35" /product="VP35"	
misc_signal	3032..3043 /gene="VP35"	
CDS	3129..4151 /note="putative; transcription start signal" /gene="VP35" /function="polymerase complex protein" /codon_start=1 /product="VP35" /protein_id="AAQ55046.1" /db_xref="GI:33860542"	
	translation="MTTRTKGRGHTAATTQNDMPGPELSGWI SEOLMTGRI PVSDLIFCDENNPGLCYASQOQTKPRTNRSQQTQDPI CNHSEFVQTLTASLTVVQQQIIHASESLQRITSLSENGUKPYVDMAKTISSLNRPNCAEMVAKYDILLVMTTGRATATAAATIEAWAEHQWPPGSLPSESAIKGIESRDETVPQSRFNNLDSTSLTSEENFGKQDEISAKDLRNTMYDHPGLGFGTAFHOLVQVICLGDQSNLSLDIIHAEFQASLAEGSPQCA LIQITKRVFIFQDAAPFVTHIRSGDI PRACQKSLRPVPSPSPKIDRGWVCVFQLDQDGK TIGLKI"	
gene	4390..5894 /gene="VP40"	
mRNA	4390..5894	


```
QY 481 CTGTATGATCGACTTGCTTCCACAGTATTCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
DB |||||
DB 6519 CTGTATGACGACTTGCTTCCACAGTTACTACCGAGGAACGACTTTTCGCTGAAGGTGTC 6578
QY 541 GTTGCATTTCTGATCTGCCCCAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
DB |||||
DB 6579 GTTGCATTTCTGATCTGCCCCAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 6638
QY 601 GAGCCGGTCAATGCAACGAGGAGACCGTCTAGTGGCTACTATTCTACCAACAATTAGATAT 660
DB |||||
DB 6639 GAGCCGGTCAATGCAACGAGGAGACCGTCTAGTGGCTACTATTCTACCAACAATTAGATAT 6698
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCAGAGTTGACAATTGACC 720
DB |||||
DB 6699 CAAGCTACCGGTTTGGAAACCAATGAGACAGAGTATTGTTCAGAGTTGACAATTGACC 6758
QY 721 TAGCTCCAACTGGAATCAAGATTTCACACACAGTTTCTGCTCCAGTGAATGAGACAATA 780
DB |||||
DB 6759 TAGCTCCAACTGGAATCAAGATTTCACACACAGTTTCTGCTCCAGTGAATGAGACAATA 6818
QY 781 TATACAAGTGGGAAGAGGAGCAATACCGGGAAACTAATTTGGAAGGTCAACCCCGAA 840
DB |||||
DB 6819 TATACAAGTGGGAAGAGGAGCAATACCGGGAAACTAATTTGGAAGGTCAACCCCGAA 6878
QY 841 ATTGATACAACAATCGGGAGTGGGCTTCTCGGAACTTAAAAAAACCTCACTAGAAAA 900
DB |||||
DB 6879 ATTGATACAACAATCGGGAGTGGGCTTCTCGGAACTTAAAAAAACCTCACTAGAAAA 6938
QY 901 ATTCCAGTGAAGAGTTGCTTTTCAAGTTGATCAACGGAGGCCAAAACATCATGTGT 960
DB |||||
DB 6939 ATTCCAGTGAAGAGTTGCTTTTCAAGTTGATCAACGGAGGCCAAAACATCATGTGT 6998
QY 961 CAGAGTCCGGCGCGAATCTTCTCCGACCCAGGAGCCACACACAACTGAAGACACACAA 1020
DB |||||
DB 6999 CAGAGTCCGGCGCGAATCTTCTCCGACCCAGGAGCCACACAACTGAAGACACACAA 7058
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1080
DB |||||
DB 7059 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 7118
QY 1081 GCAGTGTGCGATCTAACAACCTTGGCCACAATCTCCACAGTCCCGCAATCTCTCAACCC 1140
DB |||||
DB 7119 GCAGTGTGCGATCTAACAACCTTGGCCACAATCTCCACAGTCCCGCAATCTCTCAACCC 7178
QY 1141 AAACAGGTCCGGAACAGCAACCCATATACACCCGTGTATAAATCTTCAATCTCTGAG 1200
DB |||||
DB 7179 AAACAGGTCCGGAACAGCAACCCATATACACCCGTGTATAAATCTTCAATCTCTGAG 7238
QY 1201 GCAACTCAAGTTGAAACAACATCACCGCAGACAGACAACAGCAAGCAGAGTCCGACACT 1260
DB |||||
DB 7239 GCAACTCAAGTTGAAACAACATCACCGCAGACAGACAACAGCAAGCAGAGTCCGACACT 7298
QY 1261 CCCTCTGCCACGACCGCAGCAGACCCCAAGAGAGAGAACACCAACAGCAGCAGAGC 1320
DB |||||
DB 7299 CCCCCGCCACGACCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 7358
QY 1321 ACTGACTTCTGACCCCGCCACCAACAAGTCCCGCAACCAAGTCCCGCAACCAAGTCCCG 1380
DB |||||
DB 7359 ACCGACTTCTGACCCCGCCACCAACAAGTCCCGCAACCAAGTCCCGCAACCAAGTCCCG 7418
QY 1381 AACACAACACTCATCACCAAGATACCGGAGAGAGAGTCCGACGAGCAGCGGAGCTTAGGC 1440
DB |||||
DB 7419 AACACAACACTCATCACCAAGATACCGGAGAGAGAGTCCGACGAGCAGCGGAGCTTAGGC 7478
QY 1441 TTAATTACCAATATATTGCTGAGTCCGAGGAGTCAAGGCGGAGAGAACTCGA 1500
DB |||||
DB 7479 TTAATTACCAATATATTGCTGAGTCCGAGGAGTCAAGGCGGAGAGAGTCTCGA 7538
QY 1501 CGA 1503
DB 7539 AGA 7541
```

```
RESULT 14
EVU28077
LOCUS
DEFINITION
Zaire Ebola virus strain Zaire95 virion spike glycoprotein (SP)
gene, complete cds, and small/secreted glycoprotein precursor (SGP)
gene, complete cds.
ACCESSION
U28077
VERSION
U28077.1
KEYWORDS
U28077.1
SOURCE
GI:1695251
ORGANISM
Zaire ebolavirus (ZEBOV)
Zaire ebolavirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
REFERENCE
1 (bases 1 to 2408)
Sanchez,A., Trappier,S.G., Mahy,B.W., Peters,C.J. and Nichol,S.T.
The virion glycoproteins of Ebola viruses are encoded in two
reading frames and are expressed through transcriptional editing
Proc. Natl. Acad. Sci. U.S.A. 93 (8), 3602-3607 (1996)
JOURNAL
MEDLINE
PUBMED
622982
REFERENCE
2 (bases 1 to 2408)
Sanchez,A.
Direct Submission
Submitted (31-MAY-1995) Anthony Sanchez, Special Pathogens Branch,
Division of Viral and Rickettsial Diseases, Centers for Disease
Control and Prevention, 1600 Clifton Road, Bldg. 15, Room SB611,
Atlanta, GA 30333, USA
COMMENT
On Nov 30, 1996 this sequence version replaced gi:1041220.
FEATURES
Location/Qualifiers
1..2408
/organism="Zaire ebolavirus"
/mol_type="genomic RNA"
/strain="Zaire95"
/db_xref="taxon:186538"
1..141
1..14
/feature="transcription start site"
join(142..1026,1026..2171)
/function="cell receptor recognition/binding"
/note="structural glycoprotein"
/codon_start=1
/product="virion spike glycoprotein"
/protein_id="AAB37095.1"
/db_xref="GI:1695253"
/translation="MGVTGILQLPRDRKRTSFLLWVILFORTFESIPLGVIHNSTLQ
VSDVKLCRDKLSLTNQLRSVGLNLEGVATDVPSTKRWGFRSGVPPKVVNVEAG
EWAECYNLEIKKPDGSECLPAAPDGIKPRCRYHKVSGTGCAGGAFHKEGAF
LYRLASTVIYRGTTFAEGVAVFLILPOAKKDFSSHPLEFVNATEDFSSGYITTI
RYQATGFTNTEYLFEVDNLTYVQLSESRFTFQILLQNETIYTSGKRSNTTGKLIW
VNPEDTTIGWAFWETKNTKRISSELSFTAVSNRAKNIISGSPPTSDPGTNT
TTEDHKIMASNSAMVOVHSGREAAVSHLTLLATISTSPQPTTKPGPDNSTHTP
VYKLDISATQVQEHRTDNDSTADTPPATTAAGPLKAENTNTSGTDLADPATTT
SPQNSHETAGNNTHQDTGESASSGKLGLITWTIAGVGLITGGRARREIYNAQ
PKCPNHLWTTQDEGAAGLAWIPYFGPAEGYIEGLMNDGLICGLRLANETT
QALQFLRATTELTFSILNRKADFLQRWGTCILGPDCCICPEHDMTKNITDKID
QIITDFVDKTLPDQDNDNMTGWQRQIPAGITGVTGVIIVIALEFCICKFV"
```

```
5'UTR
misc_signal
1..14
```

```
CDS
```

```
gene
```

```
CDS
```

```
142..1236
/gene="SGP"
142..1236
/gene="SGP"
/function="unknown"
/note="nonstructural protein; editing event during
transcription occurs that links the 0 and -1 frames and
leads to the formation of the spike (surface) structure on
the virion"
/codon_start=1
/product="small/secreted glycoprotein precursor"
/protein_id="AAB37094.1"
/db_xref="GI:1695252"
/translation="MGVTGILQLPRDRKRTSFLLWVILFORTFESIPLGVIHNSTLQ
VSDVKLCRDKLSLTNQLRSVGLNLEGVATDVPSTKRWGFRSGVPPKVVNVEAG
EWAECYNLEIKKPDGSECLPAAPDGIKPRCRYHKVSGTGCAGGAFHKEGAF
LYRLASTVIYRGTTFAEGVAVFLILPOAKKDFSSHPLEFVNATEDFSSGYITTI
```



```

RYOATGCTGNETEYLFEVDNLTYVOLESFETPOFLLOLNETIYTSKESNTTGKLIWK
VNPELDTTIGWAFWETKTSLEKPAVKSLSQLYQTEPKTSVVRRELLPTQGPQ
QUKTKSWIQKIPLOWFKCTVKEGKIQCRI"
142..237
/ gene="SGP"
238..1230
/ gene="SGP"
/product="small/secreted glycoprotein"
1021..1027
/ gene="SGP"
/ note="transcription editing site"
2088..2156
/ gene="GP"
2088..2156
/ gene="GP"
/ note="encodes transmembrane anchor region"
2172..2408
2398..2408
ORIGIN

Query Match      64.5%; Score 1452.6; DB 14; Length 2408;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 1478; Conservative 0; Mismatches 24; Indels 1; Gaps 1;

QY 1 ATGGGCGTTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGGAGACATCTCTTT 60
Db 142 ATGGGTGTACAGGAATATTGCGATTACCTCGTGATCGATTCAAGGAGACATCTCTTT 201
QY 61 CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
Db 202 CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 261
QY 121 AGCATTACAGGTAGTGTGACAACTAGTTTGTGTGACAACTGTGTATCCACA 180
Db 262 AGCATTACAGGTAGTGTGACAACTAGTTTGTGTGACAACTGTGTATCCACA 321
QY 181 AATCAATTGAGATCAGTTGGATCTCGAAGGGAATGGATGGCACTGACGTGCA 240
Db 322 AATCAATTGAGATCAGTTGGATCTCGAAGGGAATGGATGGCACTGACGTGCA 381
QY 241 TCTGCAACTAAAGATGGGCTTCAGTCCGCTGCCACCAAGGTGGTCAATATGAA 300
Db 382 TCTGCAACTAAAGATGGGCTTCAGTCCGCTGCCACCAAGGTGGTCAATATGAA 441
QY 301 GCTGTGAATGGCTGAAACTGTCTACATCTTGAAATCAAAAACTGACGGGAGTGAG 360
Db 442 GCTGTGAATGGCTGAAACTGTCTACATCTTGAAATCAAAAACTGACGGGAGTGAG 501
QY 361 TGCTACAGCAGCGCCAGACGGGATTCGGGCTTCCCGGTCGGGTATGTGCACAAA 420
Db 502 TGCTACAGCAGCGCCAGACGGGATTCGGGCTTCCCGGTCGGGTATGTGCACAAA 561
QY 421 GTATCAGGAACGGGACCGGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTCTTC 480
Db 562 GTATCAGGAACGGGACCGGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTCTTC 621
QY 481 CTGTATGATCGACTTCTTCCAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
Db 622 CTGTATGATCGACTTCTTCCAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 681
QY 541 GTTGATTTCTGATCTGCCCCCAAGCTAAGAGGACTTCTTCAGTCTACACCCCTTGAGA 600
Db 682 GTTGATTTCTGATCTGCCCCCAAGCTAAGAGGACTTCTTCAGTCTACACCCCTTGAGA 741
QY 601 GAGCCGGTCAATGCAACGAGGACCCGCTAGTGCTACTATTCTACCAACAATTAGATAT 660
Db 742 GAGCCGGTCAATGCAACGAGGACCCGCTAGTGCTACTATTCTACCAACAATTAGATAT 801
QY 661 CAGGCTACCGGTTTTGGAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
Db 802 CAACTACCGGTTTTGGAACCAATGAGACAGAGTATTGTTGAGGTTGACAATTTGACC 861
QY 721 TACGTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780

```

```

Db 862 TACGTCCAACTTGAATCAAGATTTCACACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 921
QY 781 TATACAAGTGGGAAAAGAGGACCAATACCAACGGGAAAACCTAATTTTGAAGGTCAACCCCGAA 840
Db 922 TATACAAGTGGGAAAAGAGGACCAATACCAACGGGAAAACCTAATTTTGAAGGTCAACCCCGAA 981
QY 841 ATTGATACAACTATCGGGAGTGGGCTTCTGGAAGAACTAAAAAACCTCACTAGAAAA 900
Db 982 ATTGATACAACTATCGGGAGTGGGCTTCTGGAAGAACT-AAAAAACCTCACTAGAAAA 1040
QY 901 ATTCGAGTGAAGAGTTGTCTTTTTCACAGTTGTATCAACAGGAGCCAAAAACATCAGTGT 960
Db 1041 ATTCGAGTGAAGAGTTGTCTTTTTCACAGTGTATCAACAGGAGCCAAAAACATCAGTGT 1100
QY 961 CAGAGTCCGGCGCGAACTTCTTCGAGCCCGAGGACCAACAACTCTGAAGACACAAA 1020
Db 1101 CAGAGTCCGGCGCGAACTTCTTCGAGCCCGAGGACCAACAACTCTGAAGACACAAA 1160
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
Db 1161 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1220
QY 1081 GCAGTGTGCACTTAAACAACCTTTGCCACAATCTCCACAGTCCCAATCCTCAACAACC 1140
Db 1221 GCAGTGTGCACTTAAACAACCTTTGCCACAATCTCCACAGTCCCTCAACCCCCCAACC 1280
QY 1141 AAACCAAGTCCGGACCAACAGCACCCCAATAATACCCCGTGTATAAACTTTGACATCTCTGAG 1200
Db 1281 AAACCAAGTCCGGACCAACAGCACCCCAATAATACCCCGTGTATAAACTTTGACATCTCTGAG 1340
QY 1201 GCAACTCAAGTTGAACAACTACCGCGAGAACAGACAAACAGCAGACAGCTTCGACACT 1260
Db 1341 GCAACTCAAGTTGAACAACTACCGCGAGAACAGACAAACAGCAGACAGCTTCGACACT 1400
QY 1261 CCCTCTGCACAGCGGACCGGACCCCAAAAGCAGAGAACCAACACAGCAAGAGC 1320
Db 1401 CCCCCCGCACGACCGGACCGGACCCCTAAAGAGAGAACCAACACAGCAAGAGGT 1460
QY 1321 ACTGACTTCTGAGCCCGCCACCAACAACTAGTCCCAAAACCAACAGCAGACCGCTGCGC 1380
Db 1461 ACCGACTCTCTGAGCCCGCCACCAACAACTAGTCCCAAAACCAACAGCAGACCGCTGCGC 1520
QY 1381 AACCAACACTCATCAACAAAGATAACGGAGAGAGAGTGCAGAGCGGGAAGTACG 1440
Db 1521 AACCAACACTCATCAACAAAGATAACGGAGAGAGAGTGCAGAGCGGGAAGTACG 1580
QY 1441 TTAATTACCACTATTCTGAGTGCAGGACTGATCAAGGGGGAGAGAACTCGA 1500
Db 1581 TTAATTACCACTATTCTGAGTGCAGGACTGATCAAGGGGGAGAGAACTCGA 1640
QY 1501 CGA 1503
Db 1641 AGA 1643

RESULT 15
EVU77384
LOCUS
DEFINITION
Zaire Ebola virus strain Gabon-94 virion spike glycoprotein (GP)
U77384
ACCESSION
U77384
VERSION
U77384.1
KEYWORDS
GI:2138276
ORGANISM
Zaire ebolavirus (ZEBOV)
Zaire ebolavirus
VIRUSES; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Ebola-like viruses.
1 (bases 1 to 2446)
AUTHORS
Volchkov,V.E., Blinov,V.M., Kotov,A., Chepurnov,A. and Netesov,S.
TITLE
The full-length nucleotide sequence of the Ebola virus
JOURNAL
Thesis (1993) Institut f. Virologie, Philipps-Universitaet
REFERENCE
2 (bases 1 to 2446)

```


Qy	1021	ATCATGGCTTCAGAAAATTCCTCTCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT	1080
Db	1161	ATCATGGCTTCAGAAAATTCCTCTCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT	1220
Qy	1081	GCAGTGTGCGCATCTAACCAACCCCTTGCCCAATCTCCAGGAGTCCCAATCCCTCACAACC	1140
Db	1221	GCAGTGTGCGCATCTGACAACCCCTTGCCCAATCTCCAGGAGTCTTCGACCCCCCATACC	1280
Qy	1141	AAACGAGTCCGGACAACAGCAACCCATATATACCCGTTGTATATAACTTTGACATCTCTGAG	1200
Db	1281	AAACGAGTCCGGACAACAGCAACCCCAATATACCCGTTGTATATAACTTTGACATCTCTGAG	1340
Qy	1201	GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAACGACAGCACAGCCTCCGACACT	1260
Db	1341	GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAACGACAGCACAACTCCGACACT	1400
Qy	1261	CCCTCTGCCACGACCGCGACCGGACCCCAAGCAGAGAACCAACACGAGCAAGAGC	1320
Db	1401	CCCCCGCCACGACCGCGACCGGACCCCTTAAAGCAGAGAACCAACACGAGCAAGAGC	1460
Qy	1321	ACTGACTTCTGGACCCCGCGACCAACAAGTCCCAAAACCAAGCGAGACCGCTGGC	1380
Db	1461	ACTGACTTCTGGACCCCGCGACCAACAAGTCCCAAAACCAAGCGAGACCGCTGGC	1520
Qy	1381	AACAACACACTCATCACCAAGATACCGGAGAGAGTGCCAGCAGCGGAGGCTAGGC	1440
Db	1521	AACAACACACTCATCACCAAGATACCGGAGAGAGTGCCAGCAGCGGAGGCTAGGC	1580
Qy	1441	TTAATTACCAATPACTATTGCTGGAGTCGAGGACTGATCACAGCGGGAGAGAACTCGA	1500
Db	1581	TTAATTACCAATPACTATTGCTGGAGTCGAGGACTGATCACAGCGGGAGAGAACTCGA	1640
Qy	1501	CGA 1503	
Db	1641	AGA 1643	

Search completed: August 1, 2005, 02:43:36
Job time : 9684.51 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:55:23 ; Search time 1179.79 Seconds
(without alignments)
11299.693 Million cell updates/sec

Title: US-10-066-506A-1
Perfect score: 2252
Sequence: 1 atgggcgttacaggaatatt.....tatatcgataacggaattc 2252

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues 8780412

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: N Geneseq 16Dec04:*
 - 1: Geneseqn1980s:*
 - 2: Geneseqn1990s:*
 - 3: Geneseqn2000s:*
 - 4: Geneseqn2001as:*
 - 5: Geneseqn2001bs:*
 - 6: Geneseqn2002as:*
 - 7: Geneseqn2002bs:*
 - 8: Geneseqn2003as:*
 - 9: Geneseqn2003bs:*
 - 10: Geneseqn2003cs:*
 - 11: Geneseqn2003ds:*
 - 12: Geneseqn2004as:*
 - 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2252	100.0	2252	10	Abt13450 DNA encod
2	1506	66.9	2039	10	Abt13455 DNA encod
3	1501.4	66.7	2172	4	Aad04041 Ebola vir
4	1501.4	66.7	2224	3	Aaz51038 Ebola vir
5	1501.4	66.7	2298	3	Aaz87189 Ebola vir
6	1501.4	66.7	2298	4	Aaf76953 Ebola vir
7	1501.4	66.7	2298	12	Adl27460 Nucleotid
8	1501.4	66.7	6887	10	Acc71526 VRC6005 (
9	1501.4	66.7	6914	10	Acc71529 VRC6008 (
10	1501.4	66.7	7044	10	Acc71527 VRC6006 (
11	1501.4	66.7	7106	10	Acc71528 VRC6007 (
12	1501.4	66.7	7154	10	Acc71521 VRC6000 (
13	1501.4	66.7	7188	10	Acc71522 VRC6001 (
14	1501.4	66.7	8199	10	Acc71549 VRC6604 (
15	1501.4	66.7	8439	10	Acc71548 VRC6603 (
16	1501.4	66.7	10783	10	Acc71545 VRC6600 (
17	1499.8	66.6	7285	12	Aax59392 Plasmid p
18	1499.8	66.6	9106	12	Adf85419 Ebola vir
19	1495	66.4	7285	2	Aax89797 DNA of pv
20	1489.4	66.1	18959	13	Adm48337 Zaire ebo

21	1489.4	66.1	18959	13	Adm48364	Adm48364 Zaire ebo
22	1489.4	66.1	18959	13	Adm48355	Adm48355 Zaire ebo
23	1487.8	66.1	7272	2	Aax59393	Aax59393 Plasmid p
24	1484.6	65.9	7272	2	Aax89798	Aax89798 DNA of pv
25	1480	65.7	6724	10	Acc71525	Acc71525 VRC6004 (
26	906.4	40.2	6624	10	Acc71523	Acc71523 VRC6002 (
27	905.4	40.2	6561	10	Acc71524	Acc71524 VRC6003 (
28	900	40.0	6324	10	Acc71557	Acc71557 CMV/R Ebo
29	753	33.4	2051	10	Abt13454	Abt13454 DNA encod
30	746.6	33.2	2046	10	Abt13453	Abt13453 DNA encod
31	742.2	33.0	2100	3	Aaz51039	Aaz51039 Marburg v
32	742.2	33.0	2103	8	Acc57764	Acc57764 Marburg v
33	742.2	33.0	7778	10	Acc71550	Acc71550 VRC6701 (
34	742.2	33.0	11460	3	Aaz87211	Aaz87211 VEE repli
35	740.2	32.9	2164	4	Aad04042	Aad04042 Marburg v
36	685.8	30.5	19112	13	Adm48345	Adm48345 Marburg v
37	675.2	30.0	6467	10	Acc71530	Acc71530 VRC6052 (
38	644.2	28.6	7005	10	Acc71551	Acc71551 VRC6702 (
39	644.2	28.6	8256	10	Acc71552	Acc71552 VRC6710 (
40	548.2	24.3	2046	10	Abt13456	Abt13456 DNA encod
41	546.6	24.3	2046	10	Abt13452	Abt13452 DNA encod
42	546.2	24.3	6885	10	Acc71538	Acc71538 VRC6302 (
43	546.2	24.3	7002	10	Acc71536	Acc71536 VRC6300 (
44	546.2	24.3	7023	10	Acc71541	Acc71541 VRC6351 (
45	546.2	24.3	7036	10	Acc71537	Acc71537 VRC6301 (

ALIGNMENTS

RESULT 1
ABT13450
ID ABT13450 standard; DNA; 2252 BP.
XX
AC ABT13450;
XX
DT 30-JAN-2003 (first entry)
XX
DE DNA encoding a chimeric filovirus protein EBOV-GP1/MGV-GP2.
XX
KW Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
PN WO200279239-A2.
XX
PD 10-OCT-2002.
XX
PF 31-JAN-2002; 2002WO-US003339.
XX
PR 31-JAN-2001; 2001US-0267522P.
XX
PA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX
PI Grogan CC, Hevey MC, Schmaljohn AL;
XX
DR WPI; 2003-040651/03.
DR P-PSDB; ABJ18472.
XX
PT New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
XX
PS Claim 12; Page 66-68; 94pp; English.
XX
CC The invention relates to a chimeric filovirus glycoprotein (GP) protein
CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
CC of GP2. The chimeric filovirus GP protein is useful for inducing an
CC immune response against infection of different filoviruses, specifically
CC against both Ebola and Marburg viruses by being used as a vaccine. This
CC polynucleotide sequence represents a DNA encoding a chimeric filovirus

CC	protein of the invention
XX	Sequence 2252 BP; 679 A; 533 C; 505 G; 535 T; 0 U; 0 Other;
SQ	Query Match 100.0%; Score 2252; DB 10; Length 2252; Best Local Similarity 100.0%; Pred. No. 0; Matches 2252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGGCGGTTACAGGAATATTGCAGTATCTCAGTGATCGATTCAGAGGACATCATTTCTTT 60
DB	1 ATGGCGGTACAGGAATATTGCAGTACCCTCGTGATCGATTCAGAGGACATCATTTCTTT 60
QY	61 CTTTGGGTAAATATCTTTTCCAAGAACATTTTTCCATGCCATCTTGAGTCATCCACAAT 120
DB	61 CTTTGGGTAAATATCTTTTCCAAGAACATTTTTCCATGCCATCTTGAGTCATCCACAAT 120
QY	121 AGCACATTACAGGTTAGTGATGTCGACAAACTAGTGTTCGTGCAAACTGTCTATCCACA 180
DB	121 AGCACATTACAGGTTAGTGATGTCGACAAACTAGTGTTCGTGCAAACTGTCTATCCACA 180
QY	181 AATCAAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGGAGTGGCAA CTGACGTGCCA 240
DB	181 AATCAAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGGAGTGGCAA CTGACGTGCCA 240
QY	241 TCTGCAACTAAAAGATGGGGCTTCAGGTCGGGTCTCCCAAAAGGTGGTCAATTAATGAA 300
DB	241 TCTGCAACTAAAAGATGGGGCTTCAGGTCGGGTCTCCCAAAAGGTGGTCAATTAATGAA 300
QY	301 GCTGGTGAATGGGCTGAAAACTGCTACAACTTTGAAATCAAAAACTGACGGGAGTGAG 360
DB	301 GCTGGTGAATGGGCTGAAAACTGCTACAACTTTGAAATCAAAAACTGACGGGAGTGAG 360
QY	361 TGTCTACAGCAGCGCCAGCGGATTCGGGGCTTCCC CGGTATGTGCACAAA 420
DB	361 TGTCTACAGCAGCGCCAGCGGATTCGGGGCTTCCC CGGTATGTGCACAAA 420
QY	421 GTATCAGAACCGGACCGGTGTCGGAGACTTTGCTTCCATAAAGAGGGTGCTTTCTTC 480
DB	421 GTATCAGAACCGGACCGGTGTCGGAGACTTTGCTTCCATAAAGAGGGTGCTTTCTTC 480
QY	481 CTGTATGATCGACTTGTCTCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
DB	481 CTGTATGATCGACTTGTCTCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
QY	541 GTTCGATTTCTGATPACTGCCCAAGCTTAAGAAGACTTTTCAGCTCACACCCCTTGAGA 600
DB	541 GTTCGATTTCTGATPACTGCCCAAGCTTAAGAAGACTTTTCAGCTCACACCCCTTGAGA 600
QY	601 GAGCCGGTCAA TGCAA CGSAGGACCCGCTAGTGGCTACTATTCTACCA CAATTAGATAT 660
DB	601 GAGCCGGTCAA TGCAA CGSAGGACCCGCTAGTGGCTACTATTCTACCA CAATTAGATAT 660
QY	661 CAGGCTACCGGTTTTGGNACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
DB	661 CAGGCTACCGGTTTTGGNACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
QY	721 TAGCTFCAAATTTGAATCAAGATTCACACCAAGTTTCTGCTCCAGCTGGAATGAGACAATA 780
DB	721 TAGCTFCAAATTTGAATCAAGATTCACACCAAGTTTCTGCTCCAGCTGGAATGAGACAATA 780
QY	781 TATACAAGTGGGAAAAGAGGACNAATACCA CGGAAAACTAATTTTGGAAAGGTCAACCCCGAA 840
DB	781 TATACAAGTGGGAAAAGAGGACNAATACCA CGGAAAACTAATTTTGGAAAGGTCAACCCCGAA 840
QY	841 ATTGATACAA CAATCGGGGAGTGGGCTTTCTGGGAAA ACTAAAAAACCCTCATAGAAAA 900
DB	841 ATTGATACAA CAATCGGGGAGTGGGCTTTCTGGGAAA ACTAAAAAACCCTCATAGAAAA 900
QY	901 ATTCCAGTGAAGAGTGTCTTTTCCAGTGTCTATCAAACGGGACCAAAAAACATCAGTGET 960
DB	901 ATTCCAGTGAAGAGTGTCTTTTCCAGTGTCTATCAAACGGGACCAAAAAACATCAGTGET 960
QY	961 CAGAGTCCGGCGGCAACTTTCTTCGACCCGAGGACCAACCAAA CNACTGAAGACCAAAA 1020
DB	961 CAGAGTCCGGCGGCAACTTTCTTCGACCCGAGGACCAACCAAA CNACTGAAGACCAAAA 1020

```
Db 2041 GACTTGTCAAAATAATTTTCAGAGCAAAATTCAGCAAAATTTAAAAGGACGACAAAAGAG 2100
Qy 2101 GGGACTGTTGGGGTCTGGGTGTAATGGTGGACATCCGACTGGGGTGTCTTACTAAC 2160
Db 2101 GGGACTGTTGGGGTCTGGGTGTAATGGTGGACATCCGACTGGGGTGTCTTACTAAC 2160
Qy 2161 TTGGGCATTTTGCTACTATATTCATATGCTCTCTGATTTGCTCTATCTCTGATTTGTCGT 2220
Db 2161 TTGGGCATTTTGCTACTATATTCATATGCTCTCTGATTTGCTCTATCTCTGATTTGTCGT 2220
Qy 2221 ATCTTTACTAAATATATCGGATAACGGAAATTC 2252
Db 2221 ATCTTTACTAAATATATCGGATAACGGAAATTC 2252

RESULT 2
AB113455
ID AB113455 standard; DNA; 2039 BP.
XX
AC AB113455;
XX
DT 30-JAN-2003 (first entry)
XX
DE DNA encoding a chimeric filovirus protein EBOV-GP1/GP2.
XX
KW Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
PN WO200279239-A2.
XX
PD 10-OCT-2002.
XX
PF 31-JAN-2002; 2002WO-US003339.
XX
PR 31-JAN-2001; 2001US-0267522P.
XX
PA (USSA ) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX
PI Grogan CC, Hevey MC, Schmaljohn AL;
XX
DR WPI; 2003-040651/03.
DR P-PSDB; ABJ18477.
XX
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
XX
PS Disclosure; Page 85-88; 94pp; English.
XX
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
CC of GP2. The chimeric filovirus GP protein is useful for inducing an
CC immune response against infection of different filoviruses, specifically
CC against both Ebola and Marburg viruses by being used as a vaccine. This
CC polynucleotide sequence represents a DNA encoding a chimeric filovirus
CC protein of the invention
XX
SQ Sequence 2039 BP; 613 A; 509 C; 466 G; 451 T; 0 U; 0 Other;

Query Match 66.9%; Score 1506; DB 10; Length 2039;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGTTACAGGAATATTGAGTTACCTGTGATCGATTCAAGAGGACATCATCTTTT 60
Db 1 ATGGGCGTTACAGGAATATTGAGTTACCTGTGATCGATTCAAGAGGACATCATCTTTT 60
Qy 61 CTTTGGGTAATTTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCAAAT 120
Db 61 CTTTGGGTAATTTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCAAAT 120
```

```
Qy 121 AGCACATTACAGGTTAGTGATGTCGACAAACTAGTTTGTGTCGACAAACTGTGTCATCCACA 180
Db 121 AGCACATTACAGGTTAGTGATGTCGACAAACTAGTTTGTGTCGACAAACTGTGTCATCCACA 180
Qy 181 AATCAATTGAGATCAGTTTGACTGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 240
Db 181 AATCAATTGAGATCAGTTTGACTGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 240
Qy 241 TCTGCAACTAAAAGATGGGGCTTTCAGGTCCGGTGTCCCAACAAAGGTGGTCAATATGAA 300
Db 241 TCTGCAACTAAAAGATGGGGCTTTCAGGTCCGGTGTCCCAACAAAGGTGGTCAATATGAA 300
Qy 301 GCTGCTGAATGGGCTGAAAACTGCTACAATCTTGAAATCAAAAACTGACGGAGTGAG 360
Db 301 GCTGCTGAATGGGCTGAAAACTGCTACAATCTTGAAATCAAAAACTGACGGAGTGAG 360
Qy 361 TGTCTACACAGCGCCAGACGGGATTCGGGGCTTCCCGGTGTCGGTATGTGCACAAA 420
Db 361 TGTCTACACAGCGCCAGACGGGATTCGGGGCTTCCCGGTGTCGGTATGTGCACAAA 420
Qy 421 GTATCAGGAACGGGACCGGTGTGCCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 421 GTATCAGGAACGGGACCGGTGTGCCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTGCTGTAAGGTGTC 540
Db 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTGCTGTAAGGTGTC 540
Qy 541 GTTGCAATTTCTGATACCTGCCCAAGCTTAAGAAGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 541 GTTGCAATTTCTGATACCTGCCCAAGCTTAAGAAGACTTCTTCAGCTCACACCCCTTGAGA 600
Qy 601 GAGCCGGTCAATGCAACGGAGGACCCGTCTAGTGGCTACTATTCTACCAATTAGATAT 660
Db 601 GAGCCGGTCAATGCAACGGAGGACCCGTCTAGTGGCTACTATTCTACCAATTAGATAT 660
Qy 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGATCTTGTTCGAGGTTGACAAATTTGACC 720
Db 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGATCTTGTTCGAGGTTGACAAATTTGACC 720
Qy 721 TACGTCCAACTTGAATCAAGATTCACACCAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 721 TACGTCCAACTTGAATCAAGATTCACACCAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Qy 781 TATACAAAGTGGGAAAGGAGCAATACCAACGGGAAACTTAATTTGGAAGGTCAACCCGAA 840
Db 781 TATACAAAGTGGGAAAGGAGCAATACCAACGGGAAACTTAATTTGGAAGGTCAACCCGAA 840
Qy 841 ATTGATACAAATCGGGGAGTGGGCTTCTGGGAAACTTAAAAAACCCTCACTAGAAAA 900
Db 841 ATTGATACAAATCGGGGAGTGGGCTTCTGGGAAACTTAAAAAACCCTCACTAGAAAA 900
Qy 901 ATTCGCAAGTGAAGAGTTGCTTTTTCACAGTTGTAATCAACCGGACCAAAAAATCATCAGTGT 960
Db 901 ATTCGCAAGTGAAGAGTTGCTTTTTCACAGTTGTAATCAACCGGACCAAAAAATCATCAGTGT 960
Qy 961 CAGAGTCCGGCGCAACTTCTTCGACCCAGGGACCAACAACTGAAGAGACCAAAA 1020
Db 961 CAGAGTCCGGCGCAACTTCTTCGACCCAGGGACCAACAACTGAAGAGACCAAAA 1020
Qy 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCAAGTCAAGGGAAGGAAGCT 1080
Db 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCAAGTCAAGGGAAGGAAGCT 1080
Qy 1081 GCAGTGTGCAATCTAACCAACCTTGGCCACATCTCCAGGTCCTCCCAATCCCTCACAACC 1140
Db 1081 GCAGTGTGCAATCTAACCAACCTTGGCCACATCTCCAGGTCCTCCCAATCCCTCACAACC 1140
Qy 1141 AAACAGGTCCGGGACCAACAGCACCATAATACACCCGTGTATAAACTTGACATCTCTGAG 1200
Db 1141 AAACAGGTCCGGGACCAACAGCACCATAATACACCCGTGTATAAACTTGACATCTCTGAG 1200
```

QY 1201 GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAACAGACAGACAGCTCCGACACT 1260
Db |||||
QY 1201 GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAACAGACAGACAGCTCCGACACT 1260
Db |||||
QY 1261 CCCTCTGCCACGACCGCAGCGGACCCCAAGAGAGAGAGAACCAACACGAGCAAGAGC 1320
Db |||||
QY 1261 CCCTCTGCCACGACCGCAGCGGACCCCAAGAGAGAGAGAACCAACACGAGCAAGAGC 1320
Db |||||
QY 1321 ACTGACTTCTCTGACCCCGCCACCAACAAGTCCCAAAACACAGCAGACCGCTGGC 1380
Db |||||
QY 1321 ACTGACTTCTCTGACCCCGCCACCAACAAGTCCCAAAACACAGCAGACCGCTGGC 1380
Db |||||
QY 1381 AACAAACAACATCATCAACAAGATACCGGAGAGAGTGCCAGCAGCGGGAAGCTAGGC 1440
Db |||||
QY 1381 AACAAACAACATCATCAACAAGATACCGGAGAGAGTGCCAGCAGCGGGAAGCTAGGC 1440
Db |||||
QY 1441 TTAATTACCAATACTATTGCTGGAGTCGAGACTGATCAACAGCGGGAGAGAACTCGA 1500
Db |||||
QY 1441 TTAATTACCAATACTATTGCTGGAGTCGAGACTGATCAACAGCGGGAGAGAACTCGA 1500
Db |||||
QY 1501 CGATCG 1506
Db |||||
QY 1501 CGATCG 1506
Db |||||

RESULT 3

AAD04041
ID AAD04041 standard; DNA; 2172 BP.
XX
AC AAD04041;
XX
DT 02-JUL-2001 (first entry)
XX
DE Ebola virus strain Zaire 95 glycoprotein (GP) DNA.
XX
KW Glycoprotein; GP; immune response; vaccine; antiviral; ds.
XX
OS Ebola virus.
XX
FH Key Location/Qualifiers
FT 142..2172
FT CDS /*tag= a
FT /product= "Ebola virus strain Zaire 95 glycoprotein (GP)"
XX
PN US6200959-B1.
XX
PD 13-MAR-2001.
XX
PF 04-DEC-1996; 96US-00760615.
XX
PR 04-DEC-1996; 96US-00760615.
XX
PA (POWDER-) POWDERJECT VACCINES INC.
XX
PI Haynes JR, Schmaljohn CS, Fuller DL, Schmaljohn A, Jahrling PB;
XX
DR WPI; 2001-280564/29.
DR P-PSDB; AAE00706.
XX

Inducing an immune response to a Marburg or Ebola virus involves delivering a genetic vaccine, which contains a genetic construct encoding antigenic determinants for filovirus, using a particle acceleration device.

Claim 4; Col 17-22; 33pp; English.

The patent discloses a method of inducing an immune response to Marburg or Ebola virus glycoprotein which involves delivering a genetic vaccine, containing a genetic construct encoding antigenic determinants for filovirus, using a particle acceleration device. The genetic vaccine for filovirus is created by joining a DNA sequence encoding at least a portion of the filovirus glycoprotein to a promoter effective to promote transcription of the DNA sequence. This method is useful for inducing

CC humoral, cell-mediated and secretory immune responses in the treated individual. The present sequence is a DNA encoding Ebola virus strain Zaire 95 glycoprotein (GP), which is terminally siated. This sequence is used in the construction of genetic vaccine against filoviruses
XX
SQ Sequence 2172 BP; 656 A; 534 C; 496 G; 486 T; 0 U; 0 Other;

Query Match 66.7%; Score 1501.4; DB 4; Length 2172;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCTTACAGGAATATTGTCAGTTACCTCGTGTGCAATTCAGAGGACATCATCTTT 60
Db 142 ATGGCGCTTACAGGAATATTGTCAGTTACCTCGTGTGCAATTCAGAGGACATCATCTTT 201
QY 61 CTTTGGGTAAATTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
Db 202 CTTTGGGTAAATTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 261
QY 121 AGCACATTACAGGTTAGTGATGTCGACAAACTAGTCTTGTGTCGACAAACTGTGTCATCCACA 180
Db 262 AGCACATTACAGGTTAGTGATGTCGACAAACTAGTCTTGTGTCGACAAACTGTGTCATCCACA 321
QY 181 AATCAATTGAGATCAGTTGGACTGAACTCTCAAGGGAATGGAGTGGCACTGACGTGCA 240
Db 322 AATCAATTGAGATCAGTTGGACTGAACTCTCAAGGGAATGGAGTGGCACTGACGTGCA 381
QY 241 TCTGCAACTAAAGATGGGCTTCAGGTCCGCTGCCCAACAAAGGTGGTCAATATGAA 300
Db 382 TCTGCAACTAAAGATGGGCTTCAGGTCCGCTGCCCAACAAAGGTGGTCAATATGAA 441
QY 301 GCTGTGAATGGGCTGAAAACCTGCTCAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
Db 442 GCTGTGAATGGGCTGAAAACCTGCTCAATCTTGAATCAAAAAACCTGACGGGAGTGAG 501
QY 361 TGTCTACAGCAGCAGCAGAGCGGATTCGGGGCTTCCCGGTCGGGTATGTCACAAA 420
Db 502 TGTCTACAGCAGCAGCAGAGCGGATTCGGGGCTTCCCGGTCGGGTATGTCACAAA 561
QY 421 GTATCAGGAACGGGACCGTGTCCGGAGACTTTGCTTCCATAAAGAGGCTCTTCTTC 480
Db 562 GTATCAGGAACGGGACCGTGTCCGGAGACTTTGCTTCCATAAAGAGGCTCTTCTTC 621
QY 481 CTGTATGATCGACTTTGCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
Db 622 CTGTATGATCGACTTTGCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 681
QY 541 GTTGCAATTTCTGATCTGCTCCCAAGCTAAGAAAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 682 GTTGCAATTTCTGATCTGCTCCCAAGCTAAGAAAGGACTTCTTCAGCTCACACCCCTTGAGA 741
QY 601 GAGCCGGTCAATGCAACGGAGGACCCGCTCTAGTGGCTACTATTCTACCAATTTAGATAT 660
Db 742 GAGCCGGTCAATGCAACGGAGGACCCGCTCTAGTGGCTACTATTCTACCAATTTAGATAT 801
QY 661 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTTGTTCCAGGTTTGACAAATTTGACC 720
Db 802 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTTGTTCCAGGTTTGACAAATTTGACC 861
QY 721 TACGTCGAATTTGAATCAAGATTTCACACAGAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 862 TACGTCGAATTTGAATCAAGATTTCACACAGAGTTTCTGCTCCAGCTGAATGAGACAATA 921
QY 781 TATACAAGTGGGAAAGGAGCAATACCAAGGAAACTAATTTTGAAGGTCAACCCCGAA 840
Db 922 TATACAAGTGGGAAAGGAGCAATACCAAGGAAACTAATTTTGAAGGTCAACCCCGAA 981
QY 841 ATTGATACAAACAATCGGGAGTGGGCTTCTGGGAACTTAAAAAAACCTCTACAGAAA 900
Db 982 ATTGATACAAACAATCGGGAGTGGGCTTCTGGGAACTTAAAAAAACCTCTACAGAAA 1041
QY 901 ATTCCAGTGAAGATTGCTTTTCAGTTGATCAACAGCGGACCAAAACATCATGCTGT 960
Db |||||


```
Db 1042 ATTCCAGTGAAGATTGCTTTTCACAGTTGTATCAACAGGAGCCAAAAACATCAGTGGT 1101
Qy 961 CAGAGTCGGCGCGAATCTTTCGAGCCCGAGGACCAACAACAACACTGAAGACACAAA 1020
Db 1102 CAGAGTCGGCGCGAATCTTTCGAGCCCGAGGACCAACAACAACACTGAAGACACAAA 1161
Qy 1021 ATCATGGCTTCAGAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAGGAGCT 1080
Db 1162 ATCATGGCTTCAGAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAGGAGCT 1221
Qy 1081 GCAGTGTGCGCATCTAAACAACCTTGCACAACTCTCCAGAGTCCCAATCCCTCACAACC 1140
Db 1222 GCAGTGTGCGCATCTAAACAACCTTGCACAACTCTCCAGAGTCCCAATCCCTCACAACC 1281
Qy 1141 AAACAGTTCGGGACCAACAGCACCCATAATACACCGGTGTATAAATTGACATCTCTGAG 1200
Db 1282 AAACAGTTCGGGACCAACAGCACCCATAATACACCGGTGTATAAATTGACATCTCTGAG 1341
Qy 1201 GCAACTCAAGTTGAACAACATCACCGGAGAACAGACACACAGCAGCAGCCTCGACACT 1260
Db 1342 GCAACTCAAGTTGAACAACATCACCGGAGAACAGACACACAGCAGCAGCCTCGACACT 1401
Qy 1261 CCTCTGCGACGACGCGAGCGGACCCCAAAAGCAGAGAAACACCAACAGCAGCAAGAGC 1320
Db 1402 CCTCTGCGACGACGCGAGCGGACCCCAAAAGCAGAGAAACACCAACAGCAGCAAGAGC 1461
Qy 1321 ACTGACTTCTGGAGCCCGCCACCAACAACAAAGTCCCAAAACCAAGCAGAGACCGCTGGC 1380
Db 1462 ACTGACTTCTGGAGCCCGCCACCAACAACAAAGTCCCAAAACCAAGCAGAGACCGCTGGC 1521
Qy 1381 AACAAACAACATCATCACAGATACCGGAGAAAGAGTGCAGAGCGGGAAGCTAGGC 1440
Db 1522 AACAAACAACATCATCACAGATACCGGAGAAAGAGTGCAGAGCGGGAAGCTAGGC 1581
Qy 1441 TTAATTACCAATACATATTGCTGGAGTCCGAGACTGATCAGAGCGGGAAGAACTCGA 1500
Db 1582 TTAATTACCAATACATATTGCTGGAGTCCGAGACTGATCAGAGCGGGAAGAACTCGA 1641

Qy 1501 CGA 1503
Db 1642 AGA 1644

RESULT 4
AAZ51038
ID AAZ51038 standard; DNA; 2224 BP.
XX
AC AAZ51038;
XX
DT 05-JUN-2000 (first entry)
XX
DE Ebola virus Zaire envelope glycoprotein DNA.
XX
KW Ebola-Zaire virus envelope glycoprotein; viral glycoprotein;
KW pseudotyped retrovirus; MMLV; Moloney murine leukaemia virus; antibody;
KW transduction; screening agent; immunological agent;
KW pharmacological agent; ss.
XX
OS Ebola virus.
XX
FH Key
FT CDS 11..2041
FT /*tag= a
FT /product= "Ebola-Zaire virus envelope glycoprotein"
XX
PN WO200008131-A2.
XX
PD 17-FEB-2000.
XX
PF 04-AUG-1999; 99WO-US017702.
XX
PR 04-AUG-1998; 98US-0095242P.
PR 15-DEC-1998; 98US-0112405P.
```

```
XX (PURD ) PURDUE RES FOUND.
XX Sanders DA, Kuhn RJ, Jeffers SA, Sharkey CM, North CL;
XX Fischbach MA;
XX WPI; 2000-224030/19.
XX P-PSDB; AAY70074.
XX
XX Cells that produce inventive pseudotyped retroviruses having a broad host
XX range useful for introducing nucleotide sequences into target cells.
XX
XX Example 8; Page 58-60; 65pp; English.
XX
XX The present sequence encodes the Ebola-Zaire virus, subtype Mayinga
XX strain, envelope glycoprotein. Plasmid pEZGP1 was produced by cloning
XX into the polylinker of plasmid pcDNA3 nucleotide sequences corresponding
XX to nucleotides 6029-8253 from the Ebola Zaire virus genome. gpGFP (green
XX fluorescent protein) cells were transiently transfected with pEZGP1. This
XX is used for generation of stable cell lines transiently producing Ebola-
XX MMLV (Moloney murine leukaemia virus) pseudotyped retrovirus including
XX nucleotide sequences encoding GFP in its genome, with a broad host range.
XX These cells having different viral glycoproteins in its lipid bilayer
XX with a pseudotyped retrovirus is transduced with a desired ribonucleo-
XX side sequence. This is used to identify screening agents effective in blocking
XX viral entry into a cell. These agents may be immunological agents like
XX monoclonal or polyclonal antibodies. The pharmacological agents include
XX proteins, peptides or various chemical agents. The pseudotyped retrovirus
XX may be useful in methods of identifying cell surface receptors that allow
XX viral entry
XX
XX Sequence 2224 BP; 687 A; 541 C; 490 G; 506 T; 0 U; 0 Other;
XX
XX Query Match 66.7%; Score 1501.4; DB 3; Length 2224;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGGGTTACAGAAATATTGCAATTACCTGCTGATCGATTCAAGAGGACATCATCTTTT 60
Db 11 ATGGGGGTTACAGAAATATTGCAATTACCTGCTGATCGATTCAAGAGGACATCATCTTTT 70
Qy 61 CTTTGGGTATTTATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCCACCAAT 120
Db 71 CTTTGGGTATTTATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCCACCAAT 130
Qy 121 AGCATTATACAGTTAGTGTGCGACAAACTAGTTTGTCTGACAAACTGTCTATCCACA 180
Db 131 AGCATTATACAGTTAGTGTGCGACAAACTAGTTTGTCTGACAAACTGTCTATCCACA 190
Qy 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGAGTGCCA 240
Db 191 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGAGTGCCA 250
Qy 241 TCTGCACTAAAGATGGGGCTTCAGTCCGGTTCCTCCCAACCAAGGTGGTCAATATCAA 300
Db 251 TCTGCACTAAAGATGGGGCTTCAGTCCGGTTCCTCCCAACCAAGGTGGTCAATATCAA 310
Qy 301 GCTGTGAATGGGCTGAAAACCTGCTACAAATCTTCAAAATCAAAAACCTGACGGGAGTGAG 360
Db 311 GCTGTGAATGGGCTGAAAACCTGCTACAAATCTTCAAAATCAAAAACCTGACGGGAGTGAG 370
Qy 361 TGTCTTACAGAGCGCCAGACGGGATTCGGGGCTTCCCGGCTGCGCGGTATGTGCACAAA 420
Db 371 TGTCTTACAGAGCGCCAGACGGGATTCGGGGCTTCCCGGCTGCGCGGTATGTGCACAAA 430
Qy 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATTAAGAGGGTCTTCTTC 480
Db 431 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATTAAGAGGGTCTTCTTC 490
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
Db 491 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 550
```

```
QY 541 GTTGCAATTTCTGATCTACTGCCCCAAGCTAAGAGGACTTCTTTCAGCTCTCACACCCCTTGAGA 600
Db 551 GTTGCAATTTCTGATCTACTGCCCCAAGCTAAGAGGACTTCTTTCAGCTCTCACACCCCTTGAGA 610
QY 601 GAGCGGTCATATGCAACCGGAGGACCGCTAGTGGCTACTATTCTACCAATTTAGATAT 660
Db 611 GAGCGGTCATATGCAACCGGAGGACCGCTAGTGGCTACTATTCTACCAATTTAGATAT 670
QY 661 CAGGCTACCGGTTTGGGAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTGACC 720
Db 671 CAGGCTACCGGTTTGGGAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTGACC 730
QY 721 TAGCTCCAACTTGAATCAAGATTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 731 TAGCTCCAACTTGAATCAAGATTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 790
QY 781 TATACAGTGGGAAAGGAGCAATACCGGGGAAACTAAATTTGGAAGGTCAACCCCGAA 840
Db 791 TATACAGTGGGAAAGGAGCAATACCGGGGAAACTAAATTTGGAAGGTCAACCCCGAA 850
QY 841 ATTGATACAAATCGGGAGTGGGCTTCTGGGAAACTTAAATAAACTCACTAGAAAA 900
Db 851 ATTGATACAAATCGGGAGTGGGCTTCTGGGAACTTAAATAAACTCACTAGAAAA 910
QY 901 ATTGCAAGTGAAGAGTGTCTTTTCAAGTTGTATCAAAACGGAGCCAAAACATCAGTGGT 960
Db 911 ATTGCAAGTGAAGAGTGTCTTTTCAAGTTGTATCAAAACGGAGCCAAAACATCAGTGGT 970
QY 961 CAGAGTCCGGCGGACACTTCTCCGACCCAGGACCCACACACAACTGAGACACACAA 1020
Db 971 CAGAGTCCGGCGGACACTTCTCCGACCCAGGACCCACACAACTGAGACACACAA 1030
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1080
Db 1031 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1090
QY 1081 GCAGTGTCCGACTCTAACAAACCTTGCACAAATCTCCACGAGTCCCAATFCCCTCAACAC 1140
Db 1091 GCAGTGTCCGACTCTAACAAACCTTGCACAAATCTCCACGAGTCCCAATFCCCTCAACAC 1150
QY 1141 AAACAGGTCCGGACACAGACCCCATATACACCGGTGTATAACTTGACATCTCTGAG 1200
Db 1151 AAACAGGTCCGGACACAGACCCCATATATACACCGGTGTATAACTTGACATCTCTGAG 1210
QY 1201 GCNACTCAAGTTGAACAAATCACCGCAGACAGACAGACAGACAGACAGACAGACAGT 1260
Db 1211 GCNACTCAAGTTGAACAAATCACCGCAGACAGACAGACAGACAGACAGACAGACAGT 1270
QY 1261 CCTCTGCCACGACCGACCGGACCCCAAGAGAGAGTCCCAAAACACACGCGAGACCGCTGGC 1320
Db 1271 CCTCTGCCACGACCGACCGGACCCCAAGAGAGAGTCCCAAAACACACGCGAGACCGCTGGC 1330
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAGTCCCAAAACACACGCGAGACCGCTGGC 1380
Db 1331 ACTGACTTCTGGACCCCGCCACCAACAGTCCCAAAACACACGCGAGACCGCTGGC 1390
QY 1381 AACCAACACACTCATCAACAGATACCGGAGAGAGTCCCAAGAGAGAGTCCCAAGAGAGT 1440
Db 1391 AACCAACACACTCATCAACAGATACCGGAGAGAGTCCCAAGAGAGAGTCCCAAGAGAGT 1450
QY 1441 TTAATTTACCAATCTATTCTGGAGTCGAGGAGTCAATCAGACGGCGGGAGAGAACTCGA 1500
Db 1451 TTAATTTACCAATCTATTCTGGAGTCGAGGAGTCAATCAGACGGCGGGAGAGAACTCGA 1510
QY 1501 CGA 1503
Db 1511 AGA 1513
```

RESULT 5
AAZ87189
ID AAZ87189 standard; cDNA; 2298 BP.
XX

```
AC AAZ87189;  
XX 15-SEP-2003 (revised)  
DT 08-MAY-2000 (first entry)  
XX Ebola virus glycoprotein (GP) cDNA.  
XX Ebola virus; filovirus; glycoprotein; GP; antigen delivery;  
KW genetic vaccine; VEE virus replicon; Venezuelan equine encephalitis;  
KW Ebola haemorrhagic fever; protective immunity; diagnosis; ss.  
XX Ebola virus; 'Zaire 1976'.  
OS  
XX  
XX Key Location/Qualifiers  
FT CDS 59..2089  
FT /*tag= a  
FT /product= "Ebola virus glycoprotein (GP)"  
XX  
XX WO200000617-A2.  
XX  
XX 06-JAN-2000.  
XX  
XX 22-JUN-1999; 99WO-US014311.  
XX  
XX 29-JUN-1998; 98US-0091403P.  
XX (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.  
XX  
XX Hart MK, Wilson JA, Pushko P, Smith JP, Schmaljohn AL;  
XX WPI; 2000-160677/14.  
XX P-PSDB; AAY77112.  
XX  
XX New GP, NP, VP24, VP30, VP35 and VP40 Ebola virus proteins, useful for  
XX prevention, treatment or diagnosis of Ebola infection, particularly where  
XX expressed from virus replicons.  
XX  
XX Claim 1; Page 61; 71pp; English.  
XX The invention relates to methods and compositions for inducing a  
XX protective immune response against Ebola virus. The Ebola virus (Zaire  
XX 1976, Mayinga isolate) genes encoding glycoprotein (GP), nucleoprotein  
XX (NP), and structural proteins VP24, VP30, VP30H2, VP35 and VP40 were  
XX individually inserted into a Venezuelan equine encephalitis virus (VEE)  
XX replicon vector. The Ebola virus gene replaces the VEE virus structural  
XX protein genes. The result is a self-replicating RNA molecule that encodes  
XX its own replicase and transcriptase functions, and in addition makes  
XX abundant quantities of the Ebola virus protein. When replicon RNA is  
XX transfected into eukaryotic cells, along with two helper RNAs that  
XX express the VEE virus structural proteins, the replicon RNA is packaged  
XX into VEE virus-like particles by the VEE virus structural proteins, which  
XX are provided in trans. Since the helper RNAs lack packaging signals  
XX necessary for further propagation, the resulting VEE replicon particles  
XX (VRPs) which are produced are infectious for one cycle but are defective  
XX thereafter. On VLP infection of a cell, an abortive infection occurs  
XX whereby the cell produces the Ebola virus protein, is ultimately killed  
XX by the infection, but does not produce any viral progeny. VRPs containing  
XX a gene encoding an Ebola virus protein were injected subcutaneously into  
XX mice. VRPs expressing Ebola GP, NP, VP24, VP30, VP30H2, VP35 or VP40  
XX genes induced protective immunity in mice and may provide protection in  
XX humans. The Ebola virus proteins, nucleic acids, replicons, vectors and  
XX constructs may be used to produce vaccines against Ebola virus infection  
XX (Ebola haemorrhagic fever) in mammals, to elicit immune responses against  
XX Ebola antigens, to confer protective immunity, and to reduce disease  
XX symptoms and reduce the severity of disease. Ebola proteins and specific  
XX antibodies may be used for the diagnosis of Ebola infection. DNA  
XX sequences encoding the Ebola virus proteins may be useful as probes and  
XX primers for diagnostic hybridisation or PCR assays for detecting Ebola  
XX virus. The present sequence represents a cDNA encoding Ebola virus  
XX glycoprotein (GP). (Updated on 15-SEP-2003 to standardise OS field)  
XX  
XX Sequence 2298 BP; 703 A; 562 C; 511 G; 522 T; 0 U; 0 Other;
```

Query Match		66.7%; Score 1501.4; DB 3; Length 2298;
Best Local Similarity		99.9%; Pred. No. 0;
Matches 1502; Conservative		0; Mismatches 1; Indels 0; Gaps 0;
QY	1	ATGGCGGTTACAGGAATATTGAGTTTACCTGCTGATCGAATTCAGAGGACATCATTTCTTT 60
Db	59	ATGGCGGTTACAGGAATATTGAGTTTACCTGCTGATCGAATTCAGAGGACATCATTTCTTT 118
QY	61	CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACAAT 120
Db	119	CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACAAT 178
QY	121	AGCACATTACAGGTTAGTGTGACAAACTAGTTTGTGTCGACAACTGTCATCCACA 180
Db	179	AGCACATTACAGGTTAGTGTGACAAACTAGTTTGTGTCGACAACTGTCATCCACA 238
QY	181	AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGGAAGTGGCAACTGACGTGCCA 240
Db	239	AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGGAAGTGGCAACTGACGTGCCA 298
QY	241	TCTGCAACTAAAGATGGGGCTTACAGTCCGGTGTCCCAACCAAGGTGTCAATTATGAA 300
Db	299	TCTGCAACTAAAGATGGGGCTTACAGTCCGGTGTCCCAACCAAGGTGTCAATTATGAA 358
QY	301	GCTGTGTAATGGCTGAAACTGCTTACAACTCTTGAATCAAAACCTGACGGAGTGCAG 360
Db	359	GCTGTGTAATGGCTGAAACTGCTTACAACTCTTGAATCAAAACCTGACGGAGTGCAG 418
QY	361	TGTTCTACAGCAGCGCCAGACGGGATTCGGGCTTCCCGGTGCGGTGTGTGACAAA 420
Db	419	TGTTCTACAGCAGCGCCAGACGGGATTCGGGCTTCCCGGTGCGGTGTGTGACAAA 478
QY	421	GTATCAGGAAACGGGACCGGTGCGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 480
Db	479	GTATCAGGAAACGGGACCGGTGCGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 538
QY	481	CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGGAACGACTTTCGCTGAAGGTCTC 540
Db	539	CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGGAACGACTTTCGCTGAAGGTCTC 598
QY	541	GTTCGATTTCTGATACTGCCCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db	599	GTTCGATTTCTGATACTGCCCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 658
QY	601	GAGCCGGTCAATGCAACCGGAGGACCCGTCTAGTGGCTACTATTTCACACAATTAGATAT 660
Db	659	GAGCCGGTCAATGCAACCGGAGGACCCGTCTAGTGGCTACTATTTCACACAATTAGATAT 718
QY	661	CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTGTTCGAGTTGACAAATTTGACC 720
Db	719	CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTGTTCGAGTTGACAAATTTGACC 778
QY	721	TACGTCCCAACTTGAATCAAGATTTCACACAGATTTCCTGCTCCAGCTGAATGAGACAATA 780
Db	779	TACGTCCCAACTTGAATCAAGATTTCACACAGATTTCCTGCTCCAGCTGAATGAGACAATA 838
QY	781	TATACAAGTGGGAAAGGAGCAATACCACGGGAAACTAATTTGGAAGGTCAACCCCGAA 840
Db	839	TATACAAGTGGGAAAGGAGCAATACCACGGGAAACTAATTTGGAAGGTCAACCCCGAA 898
QY	841	ATTGATACAACTACGGGAGTGGGCTTCTGGGAAACTAATAAAAACTCTACTAGAAAA 900
Db	899	ATTGATACAACTACGGGAGTGGGCTTCTGGGAAACTAATAAAAACTCTACTAGAAAA 958
QY	901	ATTCCAGTGAAGAGTTGTCTTTCACAGTTGTATCAACCGGACCAAAAAACATCAGTGT 960
Db	959	ATTCCAGTGAAGAGTTGTCTTTCACAGTTGTATCAACCGGACCAAAAAACATCAGTGT 1018
QY	961	CAGAGTCCGGCGGAACTTTCTTCCGACCCAGGGACCAACACAACTCTGAAGACCAAAA 1020
Db	1019	CAGAGTCCGGCGGAACTTTCTTCCGACCCAGGGACCAACACAACTCTGAAGACCAAAA 1078
QY	1021	ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
Db	1079	ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1138
QY	1081	GCAGTGTGCAATCTAAACAACCCCTTTGCCACAATCTCCACGAGTCCCAATCCCTCAACAC 1140
Db	1139	GCAGTGTGCAATCTAAACAACCCCTTTGCCACAATCTCCACGAGTCCCAATCCCTCAACAC 1198
QY	1141	AAACCAAGTCCGGACCAACAGCACCCTAATAATACACCGTGTATAAACTTGACATCTCTGAG 1200
Db	1199	AAACCAAGTCCGGACCAACAGCACCCTAATAATACACCGTGTATAAACTTGACATCTCTGAG 1258
QY	1201	GCAACTCAAGTTGAACAACATCAGCAGCAACAGCAACAGCAACAGCAACAGCAACT 1260
Db	1259	GCAACTCAAGTTGAACAACATCAGCAGCAACAGCAACAGCAACAGCAACAGCAACT 1318
QY	1261	CCCTCTGCCACGACCGGACCGGACCCCAAAAGCAGAGAAACCAACAACAGCAAGAGC 1320
Db	1319	CCCTCTGCCACGACCGGACCGGACCCCAAAAGCAGAGAAACCAACAACAGCAAGAGC 1378
QY	1321	ACTGACTTCTGGACCCCGCCACCAACAAGTCCCAAAACCAACAGCAGGAGACCGCTGGC 1380
Db	1379	ACTGACTTCTGGACCCCGCCACCAACAAGTCCCAAAACCAACAGCAGGAGACCGCTGGC 1438
QY	1381	AACAACAACACTCATCAACAGATACCGGAGAGAGTGCAGCAGCGGAAGCTAGGC 1440
Db	1439	AACAACAACACTCATCAACAGATACCGGAGAGAGTGCAGCAGCGGAAGCTAGGC 1498
QY	1441	TTAATTACCAATACTATTGTGGAGTCCGAGGACTGATCAGAGCGGAGAGAACTCGA 1500
Db	1499	TTAATTACCAATACTATTGTGGAGTCCGAGGACTGATCAGAGCGGAGAGAACTCGA 1558
QY	1501	CGA 1503
Db	1559	AGA 1561
RESULT 6		
ID	AAF76953	standard; DNA; 2298 BP.
AC	AAF76953;	
DT	22-MAY-2001	(first entry)
DE	Ebola virus glycoprotein DNA sequence.	
KW	Ebola virus; glycoprotein; GP; immunogen; epitope; antibody; antiviral;	
OS	Ebola virus.	
PN	WO200116183-A1.	
PD	08-MAR-2001.	
PF	29-AUG-2000; 2000WO-US023790.	
PR	30-AUG-1999; 99US-0151505P.	
PA	(USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.	
PI	Hart MK, Wilson JA, Schmaljohn AL;	
DR	WPI; 2001-235096/24.	
PT	Novel monoclonal antibody against epitopes on the Ebola virus	
PT	glycoprotein useful as vaccines for detection, prevention and/or	
PT	therapeutic treatment of Ebola virus infections.	
PS	Disclosure; Page 55-56; 62pp; English.	
CC	The present sequence encodes Ebola virus GP, a type I transmembrane	

CC protein. The invention describes novel Ebola GP monoclonal antibodies and
CC epitopes recognised by these antibodies. Individual antibodies or
CC mixtures of antibodies are useful for preventing Ebola virus infection in
CC a subject. They are useful for inhibiting and ameliorating symptoms of
CC Ebola virus infection. They are also useful for detecting Ebola virus in
CC a sample, by incubating the sample with the antibodies, and detecting
CC antibody-Ebola virus complexes. The antibodies are also useful for
CC treating Ebola virus infection
XX

SQ Sequence 2298 BP; 703 A; 562 C; 511 G; 522 T; 0 U; 0 Other;

Query Match 66.7%; Score 1501.4; DB 4; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGTACAGGAATATTGCAGTTACCTCGTATCGATTCAAGAGGACATCATCTTT 60
DB 59 ATGGCGGTACAGGAATATTGCAGTTACCTCGTATCGATTCAAGAGGACATCATCTTT 118
QY 61 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCACAAT 120
DB 119 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCACAAT 178
QY 121 AGCATTACAGGTTAGTATGTCGACAAACTAGTTTGTGTCGACAAACTGTGTCATCCACA 180
DB 179 AGCATTACAGGTTAGTATGTCGACAAACTAGTTTGTGTCGACAAACTGTGTCATCCACA 238
QY 181 ATCAATTCAGATCAGTTGAGTGTGATCTCGAAGGGAATGGAAGTGGCACTGACGTGCCA 240
DB 239 ATCAATTCAGATCAGTTGAGTGTGATCTCGAAGGGAATGGAAGTGGCACTGACGTGCCA 298
QY 241 TCTGCACTAAAGATGCGGCTTTCAGGTCGGGTGTCCTCCACCAAGGTGCTCAATTATGAA 300
DB 299 TCTGCACTAAAGATGCGGCTTTCAGGTCGGGTGTCCTCCACCAAGGTGCTCAATTATGAA 358
QY 301 GCTGTGAAATGGCTGAAACTGCTACAATCTTGAAATCAAAAACTTGACGCGGAGTGAG 360
DB 359 GCTGTGAAATGGCTGAAACTGCTACAATCTTGAAATCAAAAACTTGACGCGGAGTGAG 418
QY 361 TGTCTACGACGACCGGACCGGATTCGGGCTTCCCGGTCGGGTATGTCGACAAA 420
DB 419 TGTCTACGACGACCGGACCGGATTCGGGCTTCCCGGTCGGGTATGTCGACAAA 478
QY 421 GTATCAGGAACCGGACCGGTGTCGGGACTTTGCTTCCATAAAGAGGTGCTTTCTTC 480
DB 479 GTATCAGGAACCGGACCGGTGTCGGGACTTTGCTTCCATAAAGAGGTGCTTTCTTC 538
QY 481 CTGTATGATCGACTTGTCTTCCAGTTATCTACCGAGGAACGACTTTCGTGAAGGTGTC 540
DB 539 CTGTATGATCGACTTGTCTTCCAGTTATCTACCGAGGAACGACTTTCGTGAAGGTGTC 598
QY 541 GTTGCATTTCTGATCTGCTCCCAAGCTAGAGGACTTCTCAGCTCACACCCCTTGAGA 600
DB 599 GTTGCATTTCTGATCTGCTCCCAAGCTAGAGGACTTCTCAGCTCACACCCCTTGAGA 658
QY 601 GAGCCGGTCAATGCAACGAGGACCGCTCTAGTGGCTACTATTCTACCAACAATTAGATAT 660
DB 659 GAGCCGGTCAATGCAACGAGGACCGCTCTAGTGGCTACTATTCTACCAACAATTAGATAT 718
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGCAAAATTGACC 720
DB 719 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGCAAAATTGACC 778
QY 721 TAGGTCACATTTGAATCAAGATTCACACAGTTTCTGCTCAGCTGATGAGCAATATA 780
DB 779 TAGGTCACATTTGAATCAAGATTCACACAGTTTCTGCTCAGCTGATGAGCAATATA 838
QY 781 TATACAAGTGGGAAAAGGAGCAATACACGCGGAAAATAATTGGAAGGTCAACCCCGAA 840
DB 839 TATACAAGTGGGAAAAGGAGCAATACACGCGGAAAATAATTGGAAGGTCAACCCCGAA 898
QY 841 ATTGATACAACTACCGGAGTGGGCTTCTGGGAAACTTAAAAAAAACCCTCATGAAAA 900

DB 899 ATTGATACAACTCGGGAGTGGCCCTTCTGGGAAACTAAAAAACCTCACTAGAAAA 958
QY 901 ATTCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAACCGGAGCCAAAAACATCAGTGT 960
DB 959 ATTCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAACCGGAGCCAAAAACATCAGTGT 1018
QY 961 CAGAGTCCGGCGGAACTTCTTCGACCCAGGAGCAACACACACTGAAGACCACAAA 1020
DB 1019 CAGAGTCCGGCGGAACTTCTTCGACCCAGGAGCAACACACACTGAAGACCACAAA 1078
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1080
DB 1079 ATCATGGCTTCAGAAAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1138
QY 1081 GCAGTGTGCGCATCTAAACAACCTTTGCCCACTCTCCACGAGTCCCAATCCCTCAACACC 1140
DB 1139 GCAGTGTGCGCATCTAAACAACCTTTGCCCACTCTCCACGAGTCCCAATCCCTCAACACC 1198
QY 1141 AAAACGAGTCCGGACAAACAGCACCCCATTAATACACCCGTGTATAAATCTTGACATCTCTGAG 1200
DB 1199 AAAACGAGTCCGGACAAACAGCACCCCATTAATACACCCGTGTATAAATCTTGACATCTCTGAG 1258
QY 1201 GCAACTCAAGTTGACACATCATCGCAGAACACACACAGCAGCACAGCCTCCGACACT 1260
DB 1259 GCAACTCAAGTTGACACATCATCGCAGAACACACACAGCAGCACAGCCTCCGACACT 1318
QY 1261 CCCTCTGCAACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACGAGCAAGAGC 1320
DB 1319 CCCTCTGCAACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACGAGCAAGAGC 1378
QY 1321 ACTGACTTCTGGACCCCGCACCAACAAGTCCCAAAACCAAGCAGGAGACCCCTGGC 1380
DB 1379 ACTGACTTCTGGACCCCGCACCAACAAGTCCCAAAACCAAGCAGGAGACCCCTGGC 1438
QY 1381 AACACACACTCATCCAGATACCCGAGAACAGAGTCCAGCAGCGGGAAGCTAGGC 1440
DB 1439 AACACACACTCATCCAGATACCCGAGAACAGAGTCCAGCAGCGGGAAGCTAGGC 1498
QY 1441 TTAATTACCAATATCTATTGCTGGAGTCCGAGGACTGATCAGAGCGGGAGAAAGAACTCGA 1500
DB 1499 TTAATTACCAATATCTATTGCTGGAGTCCGAGGACTGATCAGAGCGGGAGAAAGAACTCGA 1558
QY 1501 CGA 1503
DB 1559 AGA 1561

RESULT 7

ADL27460

ID ADL27460 standard; DNA; 2298 BP.

XX AC ADL27460;

XX AC ADL27460;

DT 20-MAY-2004 (first entry)

DE Nucleotide sequence of an Ebola virus glycoprotein.

KW glycoprotein; EGP 6D8-1-1; EGP 13F6-1-2; EGP 13C6-1-1;

KW Ebola virus infection; gene; ss.

OS Ebola virus.

FH Key Location/Qualifiers

FT CDS 59..2089

FT /tag= a

FT /product= "glycoprotein"

XX WO2004018649-A2.

PN 04-MAR-2004.

PD 21-AUG-2003; 2003WO-US027450.

XX PF

PR 23-AUG-2002; 2002US-00226795.
PA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX Hart MK, Wilson JA;
XX WPI; 2004-226835/21.
DR P-PSDB; ADL27461.
XX
PT New isolated monoclonal antibody that binds Ebola virus GP, which
PT monoclonal antibody comprises a heavy chain variable region, useful for
PT treating or ameliorating Ebola virus infection.
XX
PS Disclosure; Page 58-59; 68pp; English.
XX
CC The specification describes monoclonal antibodies that bind Ebola virus
CC glycoprotein. The antibodies are EGP 6D8-1-1, EGP 13F6-1-2, and EGP 13C6-
CC 1-1. The monoclonal antibodies of the invention are useful for treating
CC or ameliorating an Ebola virus infection. The antibodies were found to be
CC effective when administered 2 days after challenge, after significant
CC viral replication had occurred. The present sequence encodes an Ebola
CC glycoprotein. Monoclonal antibodies of the invention recognise epitopes
CC from this protein.
XX
SQ Sequence 2298 BP; 703 A; 562 C; 511 G; 522 T; 0 U; 0 Other;

Query Match 66.7%; Score 1501.4; DB 12; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGCTTACAGGAATATTCAGTTACCTCGTGATGATTCAGAGGACATCTCTTT 60
DB |||||||
QY 59 ATGGCGCTTACAGGAATATTCAGTTACCTCGTGATGATTCAGAGGACATCTCTTT 118
DB |||||||
QY 61 CTTTGGGTAAATATCTCTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
DB |||||||
QY 119 CTTTGGGTAAATATCTCTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 178
DB |||||||
QY 121 AGCATTATACAGGTAGTGATGTCGACAACTAGTTTGTGTCGACAACTGTGTCATCCACA 180
DB |||||||
QY 179 AGCATTATACAGGTAGTGATGTCGACAACTAGTTTGTGTCGACAACTGTGTCATCCACA 238
DB |||||||
QY 181 AATCAATTTGAGTACAGTTGGATGTAATCTCGAAGGGAATGAGTGGCACTGAGTGCCA 240
DB |||||||
QY 239 AATCAATTTGAGTACAGTTGGATGTAATCTCGAAGGGAATGAGTGGCACTGAGTGCCA 298
DB |||||||
QY 241 TCTGCAACTTAAAGATGGGCTTCAAGTCCGCTGTCACCAAGGCTGGTCAATTTATGAA 300
DB |||||||
QY 299 TCTGCAACTTAAAGATGGGCTTCAAGTCCGCTGTCACCAAGGCTGGTCAATTTATGAA 358
DB |||||||
QY 301 GCTGTGTAATGGCTGGAATACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 360
DB |||||||
QY 359 GCTGTGTAATGGCTGGAATACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 418
DB |||||||
QY 361 TGTCTACAGCAGCCGACGAGCGGATTCGGGCTTCCCGGCTGCGGTATGTCACAAA 420
DB |||||||
QY 419 TGTCTACAGCAGCCGACGAGCGGATTCGGGCTTCCCGGCTGCGGTATGTCACAAA 478
DB |||||||
QY 421 GTATCAGGAAACGGGACCGTGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
DB |||||||
QY 479 GTATCAGGAAACGGGACCGTGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 538
DB |||||||
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
DB |||||||
QY 539 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 598
DB |||||||
QY 541 GTTGCAATTTCTGATCTGCTCCCAACCTAAGAAGGACTTTCTCAGCTCAACCCCTTGAGA 600
DB |||||||
QY 599 GTTGCAATTTCTGATCTGCTCCCAACCTAAGAAGGACTTTCTCAGCTCAACCCCTTGAGA 658
DB |||||||
QY 601 GAGCCGGTCAATGCAACCGAGGACCGCTAGTGGCTACTATTCTACCACAATTAGATAT 660
DB |||||||
QY 659 GAGCCGGTCAATGCAACCGAGGACCGCTAGTGGCTACTATTCTACCACAATTAGATAT 718
DB |||||||

RESULT 8
ACC71526
ID ACC71526 standard; DNA; 6887 BP.

XX ACC71526;

XX 10-JUL-2003 (first entry)

XX VRC6005 (pVR1012-GP(Z) delta GP2 delta C-term A) plasmid.

XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;

KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
OS Synthetic.
XX WO2003028632-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030251.
XX 01-OCT-2001; 2001US-0326476P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Nabel GJ, Yang Z, Sullivan N, Sanchez A;
PI WPI; 2003-371961/35.
DR
XX
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
XX Claim 1; Page 126-128; 219pp; English.
XX
XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This plasmid is derived from a pVR1012 backbone and has
CC a C-terminal deletion of Ebola Glycoprotein of Zaire Subtype (GP(2))
XX
SQ Sequence 6887 BP; 1852 A; 1719 C; 1621 G; 1695 T; 0 U; 0 Other;
Query Match 66.7%; Score 1501.4; DB 10; Length 6887;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 60
DB 1939 ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 1998
QY 61 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCATCCCACTTGGAGTCAATCACAAT 120
DB 1999 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCATCCCACTTGGAGTCAATCACAAT 2058
QY 121 AGCACAATTACAGTTAGTGTGCGAACAATAGTTTGTGCGAACAATCTGTCATCCACA 180
DB 2059 AGCACAATTACAGTTAGTGTGCGAACAATAGTTTGTGCGAACAATCTGTCATCCACA 2118
QY 181 AATCAATTGAGATCAGTTGGACTGAACTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 240
DB 2119 AATCAATTGAGATCAGTTGGACTGAACTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 2178
QY 241 TTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCCAACCAAGGTGGTCAATTATGAA 300
DB 2179 TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCCAACCAAGGTGGTCAATTATGAA 2238
QY 301 GCTGGTGAATGGGCTCAAACTGCTACAATCTTGAATCAAAAAAACCCTGACGGAGTGAG 360
DB 2239 GCTGGTGAATGGGCTCAAACTGCTACAATCTTGAATCAAAAAAACCCTGACGGAGTGAG 2298
QY 361 TGTCTACAGCAGCGCCAGACGGGATTTCGGGGCTTCCCGGTGCGGTATGTGCAAAA 420

DB 2299 TGTCTACAGCAGCGCCAGACGGGATTTCGGGGCTTCCCGCGGTATGTGCAAAA 2358
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTCCTTCCATAAAGAGGGTCTTCTTC 480
DB 2359 GTATCAGGAACGGGACCGTGTGCGGAGACTTTCCTTCCATAAAGAGGGTCTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
DB 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 2478
QY 541 GTTGCAATTTCTGATACTGCCCCCAAGCTTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
DB 2479 GTTGCAATTTCTGATACTGCCCCCAAGCTTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 2538
QY 601 GAGCCGGTCAATGCAACGGGAGACCCGCTAGTGGCTACTATTCTACCACAATTAGATAT 660
DB 2539 GAGCCGGTCAATGCAACGGGAGACCCGCTAGTGGCTACTATTCTACCACAATTAGATAT 2598
QY 661 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
DB 2599 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 2658
QY 721 TAGCTCCAATTTGAATCAAGATTCAACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
DB 2659 TAGCTCCAATTTGAATCAAGATTCAACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2718
QY 781 TATCAAGTGGGAAAGGAGCAATACACCGGAAACCTAAATTTGGAAGGTCAACCCCGAA 840
DB 2719 TATCAAGTGGGAAAGGAGCAATACACCGGAAACCTAAATTTGGAAGGTCAACCCCGAA 2778
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACTAAAAAATCTCACTAGAAAA 900
DB 2779 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAACTAAAAAATCTCACTAGAAAA 2838
QY 901 ATTGCAAGTGAAGAGTTGTCTTTTCAAGTTGATCAACCGGAGCCAAAAACATCAGTGT 960
DB 2839 ATTGCAAGTGAAGAGTTGTCTTTTCAAGTTGATCAACCGGAGCCAAAAACATCAGTGT 2898
QY 961 CAGAGTCCGGCGGCAACTTCTTCGAGCCACGGGACCAACAACTGAAGACCAACAA 1020
DB 2899 CAGAGTCCGGCGGCAACTTCTTCGAGCCACGGGACCAACAACTGAAGACCAACAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1080
DB 2959 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 3018
QY 1081 GGAGTGTGCAATCTTAACAAACCTTGGCCACATCTCCAGAGTCCCAATCCCTCAACAC 1140
DB 3019 GGAGTGTGCAATCTTAACAAACCTTGGCCACATCTCCAGAGTCCCAATCCCTCAACAC 3078
QY 1141 AAACCCAGTTCGGGCAACAGCAGACCCCATATACACCCGTGTATAAACTTGACATCTCTGAG 1200
DB 3079 AAACCCAGTTCGGGCAACAGCAGACCCCATATATACCCCGTGTATAAACTTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAAACAACTACCCGCAACAGACAAACGACAGACAGCTCCGACACT 1260
DB 3139 GCAACTCAAGTTGAAACAACTACCCGCAACAGACAAACGACAGACAGCTCCGACACT 3198
QY 1261 CCCTCTGCCACGACCGGAGCGGACCCCAAAAGCAGAGAACACCAACGAGGAGGAGC 1320
DB 3199 CCCTCTGCCACGACCGGAGCGGACCCCAAAAGCAGAGAACACCAACGAGGAGGAGC 3258
QY 1321 ACTGACTTCTCTGGACCCCGCCACCAACCAAGTCCCAAAACCAACGAGGAGACCGCTGGC 1380
DB 3259 ACTGACTTCTCTGGACCCCGCCACCAACCAAGTCCCAAAACCAACGAGGAGACCGCTGGC 3318
QY 1381 AACCAACAACTCATCATCACAGATACCGGAGAGAGAGTGCACAGCGGAGCTAGGC 1440
DB 3319 AACCAACAACTCATCATCACAGATACCGGAGAGAGAGTGCACAGCGGAGCTAGGC 3378
QY 1441 TTAATTTACCAATATCTATTGCTGGAGTCCGAGACTGATCAGAGCGGGGAGAGNACTCCA 1500

Db 3379 TTAATTACCAATACTATTGCTGGAGTCGACGAGTCTGATCAGAGGGGAGAGACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441
RESULT 9
ID ACC71529 standard; DNA; 6914 BP.
XX AC ACC71529;
XX XT 10-JUL-2003 (first entry)
XX DE VRC6008 (pVR1012-GP(Z) delta TM) plasmid.
XX KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
XX KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
XX KW viral infection; filovirus; circular; cyclic; ds.
XX OS Synthetic.
XX PN WO2003028632-A2. -
XX PD 10-APR-2003.
XX PP 24-SEP-2002; 2002WO-US030251.
XX PR 01-OCT-2001; 2001US-0326476P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX DR WPI; 2003-371961/35.
XX PT New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX PS Claim 1; Page 133-135; 219pp; English.
CC The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprising of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprising of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This plasmid is derived from a pVR1012 backbone and has
CC truncation of the TM region of Ebola Glycoprotein of Zaire Subtype
XX (GP(Z))
SQ Sequence 6914 BP; 1850 A; 1750 C; 1645 G; 1669 T; 0 U; 0 Other;
Query Match 66.7%; Score 1501.4; DB 10; Length 6914;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGGGTTACAGGAATATTGAGTTACCTCGTGTGATTCAGATTCAGAGGACATCATTTCTTT 60
Db 1939 ATGGGGGTTACAGGAATATTGAGTTACCTCGTGTGATTCAGATTCAGAGGACATCATTTCTTT 1998
QY 61 CTTTGGGTTAATTCTCTTTTCCAGAGAACATTTCCATCCCACTTGGAGTCAATCCACAT 120
|||||

Db 1999 CTTTGGGTTAATTCTCTTTTCCAGAGAACATTTTCCATCCCACTTGGAGTCAATCCACAT 2058
QY 121 AGCACATTACAGGTTAGTGTGCGACAAACTAGTTGTGTGCGACAAACTGTGTATCCACA 180
Db 2059 AGCACATTACAGGTTAGTGTGCGACAAACTAGTTGTGTGCGACAAACTGTGTATCCACA 2118
QY 181 AATCAATTTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGAGTGCCA 240
Db 2119 AATCAATTTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGAGTGCCA 2178
QY 241 TCTGCAACTAAAGATGGGCTTCAGGTCCCGTGTCCACCAAAAGGTGCTCAATTTATGAA 300
Db 2179 TCTGCAACTAAAGATGGGCTTCAGGTCCCGTGTCCACCAAAAGGTGCTCAATTTATGAA 2238
QY 301 GCTGGTGAATGGGCTGAAAACCTGCTCAATCTTGAATCAAAAACCTGAGCGGAGTGAG 360
Db 2239 GCTGGTGAATGGGCTGAAAACCTGCTCAATCTTGAATCAAAAACCTGAGCGGAGTGAG 2298
QY 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGCCTGATGTGCGACAAA 420
Db 2299 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGCCTGATGTGCGACAAA 2358
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGTGTGCGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACGAGGAAACGACTTTTCGCTGAAGGTGC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACGAGGAAACGACTTTTCGCTGAAGGTGC 2478
QY 541 GTTGCATTTCTGATCTACCTCCCAAGCTTAAGAGGACTTTTTCAGCTCACACCCCTTGAGA 600
Db 2479 GTTGCATTTCTGATCTACCTCCCAAGCTTAAGAGGACTTTTTCAGCTCACACCCCTTGAGA 2538
QY 601 GAGCGGCTCAATGCAACCGGAGGACCGTCTAGTGGCTACTATTCTACCACAATTAGATAT 660
Db 2539 GAGCGGCTCAATGCAACCGGAGGACCGTCTAGTGGCTACTATTCTACCACAATTAGATAT 2598
QY 661 CAGGTACCGGTTTGGAAACCAATGAGACAGACTTCTGTTGCGAGGTTGACAAATTTGACC 720
Db 2599 CAGGTACCGGTTTGGAAACCAATGAGACAGACTTCTGTTGCGAGGTTGACAAATTTGACC 2658
QY 721 TAGCTCCAATCTGAATCAAGATTACACCAAGTTCTGCTCCAGCTCAATGAGACATA 780
Db 2659 TAGCTCCAATCTGAATCAAGATTACACCAAGTTCTGCTCCAGCTCAATGAGACATA 2718
QY 781 TATACAAGTGGGAAAGGAGCAATACCAACCGGAAACCTAAATTTGGAAGGTCAACCCCGAA 840
Db 2719 TATACAAGTGGGAAAGGAGCAATACCAACCGGAAACCTAAATTTGGAAGGTCAACCCCGAA 2778
QY 841 ATTGATACAAACATCGGGGAGTGGGCTTTCTGGGAAACCTAAAAAACCTCACTAGAAAA 900
Db 2779 ATTGATACAAACATCGGGGAGTGGGCTTTCTGGGAAACCTAAAAAACCTCACTAGAAAA 2838
QY 901 ATTGCGAGTGAAGAGTTGTCTTTTACAGTTGTATCAACGGGAGCCCAAAACATCAGTGGT 960
Db 2839 ATTGCGAGTGAAGAGTTGTCTTTTACAGTTGTATCAACGGGAGCCCAAAACATCAGTGGT 2898
QY 961 CAGAGTCCGGCGCGAACTTTCTTCGAGCCAGGAGCCAAACAACTGAAGACCAACAAA 1020
Db 2899 CAGAGTCCGGCGCGAACTTTCTTCGAGCCAGGAGCCAAACAACTGAAGACCAACAAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGTGACAGTCAAGGAGGAAAGCT 1080
Db 2959 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGTGACAGTCAAGGAGGAAAGCT 3018
QY 1081 GCAGTGTGCAATCTAAGAACCTTTGGCCCAATCTCCACAGGTCCCAATCCCTCAACACC 1140
Db 3019 GCAGTGTGCAATCTAAGAACCTTTGGCCCAATCTCCACAGGTCCCAATCCCTCAACACC 3078
QY 1141 AAACCCAGGTCCGGCAACACAGCACCCATAATACCCCGTGTATAAATTTGACATCTCTGAG 1200
Db 3079 AAACCCAGGTCCGGCAACACAGCACCCATAATACCCCGTGTATAAATTTGACATCTCTGAG 3138

QY 1201 GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAGACAGACGCTCCGACACT 1260
DB 3139 GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAGACAGACGCTCCGACACT 3198
QY 1261 CCCTCTGCCACGACCGCAGCCGACCCCCCAAAAGACGAGAACACCAACACGAGAGAGC 1320
DB 3199 CCCTCTGCCACGACCGCAGCCGACCCCCCAAAAGACGAGAACACCAACACGAGAGAGC 3258
QY 1321 ACTGACTTCTCGACCCGCCACCAACAAGTCCCAAAACACAGCGAGACCGCTGCG 1380
DB 3259 ACTGACTTCTCGACCCGCCACCAACAAGTCCCAAAACACAGCGAGACCGCTGCG 3318
QY 1381 AACAAACAACATCATCACCAAGATACCGGAGAGAGAGTCCAGCAGCGGGAAGCTAGGC 1440
DB 3319 AACAAACAACATCATCACCAAGATACCGGAGAGAGAGTCCAGCAGCGGGAAGCTAGGC 3378
QY 1441 TTAATTACCAATCTATTCTGAGTCCGAGCTGATCACAGCGGGAAGAGAACTCGA 1500
DB 3379 TTAATTACCAATCTATTCTGAGTCCGAGCTGATCACAGCGGGAAGAGAACTCGA 3438
QY 1501 CGA 1503
DB 3439 AGA 3441

RESULT 10
ACCT1527
ID ACCT1527 standard; DNA; 7044 BP.
AC ACCT1527;
XX
XX
DT 10-JUL-2003 (first entry)
DE VRC6006 (pVR1012-GP(Z) delta GP2 delta C-term B) plasmid.
XX
XX Viricide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX
OS Synthetic.
XX
XX WO2003028632-A2.
XX
XX 10-APR-2003.
XX
XX 24-SEP-2002; 2002WO-US030251.
XX
XX 01-OCT-2001; 2001US-0326476P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX WPI; 2003-371961/35.
XX
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
XX Claim 1; Page 128-130; 219pp; English.
XX
XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACCT1521-ACCT1563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected

CC from constructs ACCT1521-ACCT1563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This plasmid is derived from a pVR1012 backbone and has
CC a C-terminal deletion of Ebola Glycoprotein of Zaire Subtype (GP(Z))
XX
SQ Sequence 7044 BP; 1892 A; 1766 C; 1656 G; 1730 T; 0 U; 0 Other;
Query Match 66.7%; Score 1501.4; DB 10; Length 7044;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGCGTTACAGGAATATTCAGTTACCTCGTATGATTCAGAGACATCATCTTTT 60
DB 1939 ATGGCGTTACAGGAATATTCAGTTACCTCGTATGATTCAGAGACATCATCTTTT 1998
QY 61 CTTTGGGTAAATATTCCTTTTCCAAAGAACATTTTCCATCCACATTTGGAGTCAATC 120
DB 1999 CTTTGGGTAAATATTCCTTTTCCAAAGAACATTTTCCATCCACATTTGGAGTCAATC 2058
QY 121 AGCATTACAGGTTAGTGATGTCGACAAATCTAGTTTCTGTCGACAAATCTGTCATCC 180
DB 2059 AGCATTACAGGTTAGTGATGTCGACAAATCTAGTTTCTGTCGACAAATCTGTCATCC 2118
QY 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCACTGACGTGCCA 240
DB 2119 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCACTGACGTGCCA 2178
QY 241 TCTGCACTAAAGATGGGCTTCAGTCCGGTGTCCACCAAGGTGGTCAATATGAA 300
DB 2179 TCTGCACTAAAGATGGGCTTCAGTCCGGTGTCCACCAAGGTGGTCAATATGAA 2238
QY 301 GCTGTGAATGGGCTGAAAACTGCTACAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
DB 2239 GCTGTGAATGGGCTGAAAACTGCTACAATCTTGAATCAAAAAACCTGACGGGAGTGAG 2298
QY 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGCTTCCCGGTGCGGTATGTGACAAA 420
DB 2299 TGTCTACAGCAGCGCCAGACGGGATTCGGGCTTCCCGGTGCGGTATGTGACAAA 2358
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTCTTC 480
DB 2359 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTCTTC 2418
QY 481 CTGTATGATCGACTTGTTCACACAGTTATCTACCGAGGAACGACTTTGCTGAAAGTGT 540
DB 2419 CTGTATGATCGACTTGTTCACACAGTTATCTACCGAGGAACGACTTTGCTGAAAGTGT 2478
QY 541 GTTGCAATTTCTGATCTACTGCCCAAGCTAAGAGGACTTCTTCAGTCAACCCCTTGAGA 600
DB 2479 GTTGCAATTTCTGATCTACTGCCCAAGCTAAGAGGACTTCTTCAGTCAACCCCTTGAGA 2538
QY 601 GAGCGGTCAATGCAACGGGAGGACCGGTCTAGTGGCTACTATTCTACCAATTAAGATAT 660
DB 2539 GAGCGGTCAATGCAACGGGAGGACCGGTCTAGTGGCTACTATTCTACCAATTAAGATAT 2598
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
DB 2599 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 2658
QY 721 TACGTCCAATCTGAATCAAGATTCACACAGTTTCTGCTCCAGTGAATGAGACAATA 780
DB 2659 TACGTCCAATCTGAATCAAGATTCACACAGTTTCTGCTCCAGTGAATGAGACAATA 2718
QY 781 TATACAAGTGGGAAAGGACCAATACCGGGAACCAATTAATTTGGAAGTCAACCCGAA 840
DB 2719 TATACAAGTGGGAAAGGACCAATACCGGGAACCAATTAATTTGGAAGTCAACCCGAA 2778
QY 841 ATTGATACAACAATCGGGAGTGGGCTTCTGGGAACTTAAAAAAAACCCTACTAGAAA 900
DB 2779 ATTGATACAACAATCGGGAGTGGGCTTCTGGGAACTTAAAAAAAACCCTACTAGAAA 2838
QY 901 ATTCCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAAAACGGGAGCCCAAAACATCTAGTGT 960


```
Db 2839 ATTCCGAGTGAAGAGTTGTTCTTTCACAGTTGTATCAACGGAGCCAAAACATCATGTGT 2898
Qy 961 CAGAGTCCGGCGGAATCTTCTCCGACCCAGGAGCCAAACAACTGAAGACACACAAA 1020
Db 2899 CAGAGTCCGGCGGAATCTTCTCCGACCCAGGAGCCAAACAACTGAAGACACACAAA 2958
Qy 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAGAGGAAGGAGCT 1080
Db 2959 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAGAGGAAGGAGCT 3018
Qy 1081 GCAGTGTGCATCTAACAACTTCCGCAATCTCCACAGTCCCAATCCCTCAACACC 1140
Db 3019 GCAGTGTGCATCTAACAACTTCCGCAATCTCCACAGTCCCAATCCCTCAACACC 3078
Qy 1141 AAACAGGTCCGGCAACAGCACCCATATACACCCGTGTATATACTTGACATCTCTGAG 1200
Db 3079 AAACAGGTCCGGCAACAGCACCCATATACACCCGTGTATATACTTGACATCTCTGAG 3138
Qy 1201 GCAACTCAAGTTGAACACATCAGCGCAGACAGACAGACAGACAGACAGACAGACAGCT 1260
Db 3139 GCAACTCAAGTTGAACACATCAGCGCAGACAGACAGACAGACAGACAGACAGACAGCT 3198
Qy 1261 CCTCTGCGACAGCGCGGAGCCGCCAAAGCAGAGAACCAACACAGCAGCAAGAGC 1320
Db 3199 CCTCTGCGACAGCGCGGAGCCGCCAAAGCAGAGAACCAACACAGCAGCAAGAGC 3258
Qy 1321 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACACAGCGGAGACCCCTGGC 1380
Db 3259 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACACAGCGGAGACCCCTGGC 3318
Qy 1381 AACAAACACATCATCAACAGATACCGGAGAGAGAGTCCAGACAGCGGAGAGCTAGGC 1440
Db 3319 AACAAACACATCATCAACAGATACCGGAGAGAGAGTCCAGACAGCGGAGAGCTAGGC 3378
Qy 1441 TTAATTACCAATATTGCTGAGTCCGAGTCCGAGACTGATCAGAGCGGGAGAGACTCGA 1500
Db 3379 TTAATTACCAATATTGCTGAGTCCGAGTCCGAGACTGATCAGAGCGGGAGAGACTCGA 3438
Qy 1501 CGA 1503
Db 3439 AGA 3441

RESULT 11
ID ACC71528 standard; DNA; 7106 BP.
AC ACC71528;
XX
XX 10-JUL-2003 (first entry)
XX
XX VRC6007 (pVR1012-GP(Z) delta GP2 delta FUS) plasmid.
XX
XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;
XX Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
XX viral infection; filovirus; circular; cyclic; ds.
XX
XX Synthetic.
XX
XX WO2003028632-A2.
XX
XX 10-APR-2003.
XX
XX 24-SEP-2002; 2002WO-US030251.
XX
XX 01-OCT-2001; 2001US-0326476P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX WPI; 2003-371961/35.
XX
XX
```

```
XX
PT New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
PS Claim 1; Page 130-132; 219pp; English.
XX
CC The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This plasmid is derived from a pVR1012 backbone and has
CC a deletion of a fusion peptide in GP2 of Ebola Glycoprotein of Zaire
CC Subtype (GP(Z))
XX
SQ Sequence 7106 BP; 1923 A; 1770 C; 1667 G; 1746 T; 0 U; 0 Other;
Query Match 66.7%; Score 1501.4; DB 10; Length 7106;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGACATCATCTTT 60
Db 1939 ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGACATCATCTTT 1998
Qy 61 CTTTGGGTAAATATCTTTTCCAAAGACATTTTCCATCCCATCTGGAGTCATCCACAA 120
Db 1999 CTTTGGGTAAATATCTTTTCCAAAGACATTTTCCATCCCATCTGGAGTCATCCACAA 2058
Qy 121 AGCATTATACAGGTAGTGATGTGACAAATAGTCTGTGACAAATCTGTATCCACA 180
Db 2059 AGCATTATACAGGTAGTGATGTGACAAATAGTCTGTGACAAATCTGTATCCACA 2118
Qy 181 AATCAATTGATCATGTTGGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCA 240
Db 2119 AATCAATTGATCATGTTGGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCA 2178
Qy 241 TCTGCACTAAAGATGGGCTTCAGGTCCGGTGTCCCAACCAAGGTGGTCAATATGAA 300
Db 2179 TCTGCACTAAAGATGGGCTTCAGGTCCGGTGTCCCAACCAAGGTGGTCAATATGAA 2238
Qy 301 GCTGTGATGGGTGAAATCTGCTACAAATCTTGAATCAAAAACCTGACGGGAGTGAG 360
Db 2239 GCTGTGATGGGTGAAATCTGCTACAAATCTTGAATCAAAAACCTGACGGGAGTGAG 2298
Qy 361 TGTCTACAGACGCCACAGACGGGATTCGGGCTTCCCGGTGCGGTATGTGCACAAA 420
Db 2299 TGTCTACAGACGCCACAGACGGGATTCGGGCTTCCCGGTGCGGTATGTGCACAAA 2358
Qy 421 GTATCAGGAACGGGACCGGTGCGCGGAGACTTTGGCTTCCATTAAGAGGGTCTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGGTGCGCGGAGACTTTGGCTTCCATTAAGAGGGTCTTCTTC 2418
Qy 481 CTGTATGATCGACTTGTCTTCACAGTTATCTACAGGAAACGACTTTCCGTGAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCACAGTTATCTACAGGAAACGACTTTCCGTGAGGTGTC 2478
Qy 541 GTTGCATTTCTGATACGTGCCCCAAAGCTAAGAGGACTTTCTTTCAGCTCAACCCCTTGAG 600
Db 2479 GTTGCATTTCTGATACGTGCCCCAAAGCTAAGAGGACTTTCTTTCAGCTCAACCCCTTGAG 2538
Qy 601 GAGCGGTCAATGCAACGGGAGACCGGTCTAGTGGCTACTATTCTACCAATATGATAT 660
XX
```

Db 2539 GAGCCGGTCAATGCAACCGGAGACCGTCTAGTGGCTACTATTCTACCAATATAGATAT 2598
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTCTGTCAGGTTGACAAATTTGACC 720
Db 2599 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTCTGTCAGGTTGACAAATTTGACC 2658
QY 721 TACGTCCTCAATTTGAATCAAGATTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2659 TACGTCCTCAATTTGAATCAAGATTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2718
QY 781 TATACAGTGGGAAAGAGGACATACCAGGGAACAACTAATTTGGAGTCAACCCCGAA 840
Db 2719 TATACAGTGGGAAAGAGGACATACCAGGGAACAACTAATTTGGAGTCAACCCCGAA 2778
QY 841 ATTGATACAACTCGGCGAGTGGGCTTCTCGGAACTTAAATAAACTACTAGAAAA 900
Db 2779 ATTGATACAACTCGGCGAGTGGGCTTCTCGGAACTTAAATAAACTACTAGAAAA 2838
QY 901 ATTGCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAACCGGACCAAAAAACATCAGTGGT 960
Db 2839 ATTGCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAACCGGACCAAAAAACATCAGTGGT 2898
QY 961 CAGAGTCCGGCGGAACTTCTTCGACCCAGGACCAACACAACTGAAGACACACAA 1020
Db 2899 CAGAGTCCGGCGGAACTTCTTCGACCCAGGACCAACACAACTGAAGACACACAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAGGAGT 1080
Db 2959 ATCATGGCTTCAGAAAAATTCCTCTGCAATGTTCAAGTGCACAGTCAAGGAGGAGT 3018
QY 1081 GCAGTGTGCGATCTTAACAAACCCCTTGCACAACTCTCCAGAGTCCCAATCCCTCAACCC 1140
Db 3019 GCAGTGTGCGATCTTAACAAACCCCTTGCACAACTCTCCAGAGTCCCAATCCCTCAACCC 3078
QY 1141 AAACCGAGTCCGACCAACAGCCCAATATACACCGGTGTATAAATTGACATCTCTGAG 1200
Db 3079 AAACCGAGTCCGACCAACAGCCCAATATACACCGGTGTATAAATTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAACTACCGCAGACAGACAGACAGACAGACAGACAGACAGT 1260
Db 3139 GCAACTCAAGTTGAACAACTACCGCAGACAGACAGACAGACAGACAGACAGACAGT 3198
QY 1261 CCCTCTGCCAGACCGCAGCCGACCCCAAAAGCAGAGAGACCAACACAGCAGCAGAGC 1320
Db 3199 CCCTCTGCCAGACCGCAGCCGACCCCAAAAGCAGAGAGACCAACACAGCAGCAGAGC 3258
QY 1321 ACTGACTTCTGACCCCGCCACCAACAAAGTCCCAAAACACAGCAGCAGCAGCAGTGGC 1380
Db 3259 ACTGACTTCTGACCCCGCCACCAACAAAGTCCCAAAACACAGCAGCAGCAGCAGTGGC 3318
QY 1381 AACCAACACACTCATCACCAGATACCGGAGAGAGAGTCCAGCAGCGGAGGAGCTAGGC 1440
Db 3319 AACCAACACACTCATCACCAGATACCGGAGAGAGAGTCCAGCAGCGGAGGAGCTAGGC 3378
QY 1441 TTAATTTACCAATATCTTCTGAGTTCGAGGAGTATCAGAGGCGGAGAGAACTCGA 1500
Db 3379 TTAATTTACCAATATCTTCTGAGTTCGAGGAGTATCAGAGGCGGAGAGAACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441

RESULT 12

ACC71521
ID ACC71521 standard; DNA; 7154 BP.

XX AC

XX ACC71521;

XX DT

XX 10-JUL-2003 (first entry)

XX DE

XX VRC6000 (pVR1012-GP(2) plasmid.

XX XX

KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
OS Synthetic.
XX
XX WO2003028632-A2.
PN
XX
XX 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030251.
XX
XX 01-OCT-2001; 2001US-0326476P.
PR
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX
XX WPI; 2003-371961/35.
DR
XX
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
XX Claim 1; Page 114-116; 219pp; English.
XX

CC The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus, or
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This plasmid expresses the Ebola Glycoprotein of Zaire
CC subtype and is derived from a pVR1012 backbone

SQ Sequence 7154 BP; 1932 A; 1784 C; 1682 G; 1756 T; 0 U; 0 Other;

Query Match 66.7%; Score 1501.4; DB 10; Length 7154;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 60

Db 1939 ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 1998

QY 61 CTTTGGGTAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAACCAAT 120

Db 1999 CTTTGGGTAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAACCAAT 2058

QY 121 ACCACATTACAGGTTAGTGTGACAACTAGTTTCTCGTGACAACTGTCATCCACA 180

Db 2059 AGCATTACAGGTTAGTGTGACAACTAGTTTCTCGTGACAACTGTCATCCACA 2118

QY 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 240

Db 2119 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 2178

QY 241 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCGCCACCAAGGTGGTCAATTATGAA 300

Db 2179 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCGCCACCAAGGTGGTCAATTATGAA 2238

QY 301 GCTGGTGAATGGCTGAAAACTGCTACAACTCTTGAATCAAAAAACCTGACGGGAGTGAG 360

Db 2239 GCTGGTGAATGGCTGAAAACTGCTACAACTCTTGAATCAAAAAACCTGACGGGAGTGAG 2298

QY 361 TGTCTACGAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGTCACAAA 420
Db 2299 TGTCTACGAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGTCACAAA 2358
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGAACGACTTTCGCTGAAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGAACGACTTTCGCTGAAGGTGTC 2478
QY 541 GTTGCATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 2479 GTTGCATTTCTGATACTGCCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 2538
QY 601 GAGCCGGTCAATGCAACGAGGAGACCGCTCTAGTGGCTACTATTCTTACCAACAATTAGATAT 660
Db 2539 GAGCCGGTCAATGCAACGAGGAGACCGCTCTAGTGGCTACTATTCTTACCAACAATTAGATAT 2598
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
Db 2599 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 2658
QY 721 TACGTCCTCAACTTGAATCAAGATTTCACACAGTCTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2659 TACGTCCTCAACTTGAATCAAGATTTCACACAGTCTTCTGCTCCAGCTGAATGAGACAATA 2718
QY 781 TATACAAGTGGGAAAGAGCAATACACCGGAAATACTAATTTGGAAGTCAACCCCGAA 840
Db 2719 TATACAAGTGGGAAAGAGCAATACACCGGAAATACTAATTTGGAAGTCAACCCCGAA 2778
QY 841 ATTGATACAACAATCGGGAGTGGGCTTCTGGGAAACTTAAAAAAAATCTCACTAGAAAA 900
Db 2779 ATTGATACAACAATCGGGAGTGGGCTTCTGGGAAACTTAAAAAAAATCTCACTAGAAAA 2838
QY 901 ATTCCAGTGAAGAGTGTCTTCTCAGTGTGTATCAACGGGACCAAAAACATCAGTGT 960
Db 2839 ATTCCAGTGAAGAGTGTCTTCTCAGTGTGTATCAACGGGACCAAAAACATCAGTGT 2898
QY 961 CAGAGTCCGGCGCAACTTCTTCGACCCAGGAGCAACAACAACACTGAGACCAACAA 1020
Db 2899 CAGAGTCCGGCGCAACTTCTTCGACCCAGGAGCAACAACAACACTGAGACCAACAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAGTGCACAGTCAAGGAGGGAAGCT 1080
Db 2959 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAGTGCACAGTCAAGGAGGGAAGCT 3018
QY 1081 GCAGTGTGCACTTAACAACCTTGGCACAATCTCCAGAGTCCCAATCCCTCAACACC 1140
Db 3019 GCAGTGTGCACTTAACAACCTTGGCACAATCTCCAGAGTCCCAATCCCTCAACACC 3078
QY 1141 AAACAGGTCGGGACACAGACCCATTAATACACCGGTGTATAAATCTGACATCTCTGAG 1200
Db 3079 AAACAGGTCGGGACACAGACCCATTAATACACCGGTGTATAAATCTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAACATCACCGGAGAACAGACAACAGACAGACAGCTTCGACACT 1260
Db 3139 GCAACTCAAGTTGAACAACATCACCGGAGAACAGACAACAGACAGACAGCTTCGACACT 3198
QY 1261 CCCTCTGCCACAGCCGAGCCGACCCCAAAAGAGAGAGAACACCAACACGAGGAGAGC 1320
Db 3199 CCCTCTGCCACAGCCGAGCCGACCCCAAAAGAGAGAGAACACCAACACGAGGAGAGC 3258
QY 1321 ACTGACTTCTCGACCCGACCAACCAAGTCCCAAAACACAGCGAGACCGCTGCG 1380
Db 3259 ACTGACTTCTCGACCCGACCAACCAAGTCCCAAAACACAGCGAGACCGCTGCG 3318
QY 1381 AACAAACAACATCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGGAAGCTAGGC 1440
Db 3319 AACAAACAACATCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGGAAGCTAGGC 3378
QY 1441 TTAATTACCAATACTATTGCTGGAGTGGAGGACTGATCATCAGGCGGGAGAGAACTCGA 1500

Db 3379 TTAATTACCAATACTATTGCTGGAGTGCAGGACTGATCATCAGGCGGGAGAGAACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441
RESULT 13
ACC71522
ID ACC71522 standard; DNA; 7188 BP.
XX
AC ACC71522;
XX
DT 10-JUL-2003 (first entry)
XX
VRC6001 (pVR1012x/s-GP(Z) plasmid.
XX
Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX
Synthetic.
XX
WO2003028632-A2.
XX
10-APR-2003.
XX
24-SEP-2002; 2002WO-US030251.
XX
01-OCT-2001; 2001US-0326476P.
XX
(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX
WPI; 2003-371961/35.
XX
New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
Claim 1; Page 116-119; 219pp; English.
XX
The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This plasmid expresses the Ebola Glycoprotein of Zaire
CC Subtype and is derived from a pVR1012 backbone
XX
SQ Sequence 7188 BP; 1942 A; 1791 C; 1689 G; 1766 T; 0 U; 0 Other;
Query Match 66.7%; Score 1501.4; DB 10; Length 7188;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGTTACAGGAATATTGCGTTACCTCGTGATCGATTCAAGAGACATCATCTTT 60
Db 1939 ATGGGCGTTACAGGAATATTGCGTTACCTCGTGATCGATTCAAGAGACATCATCTTT 1998
QY 61 CTTTGGGTAAATATCTTTTCCAAAGAACATTTTCCATCCCACTTGGATCATCCAAAT 120

1999 CTTTGGGTAAATTAATCCTTTTCCAAAGAAACATTTTCCATCCCACTTTGGAGTCAATCCACAAAT 2058
121 AGCATTACAGGTTAGTGATCGACAAACTAGTTTGTCTGTGACAAACTGTGATCCACA 180
2059 AGCATTACAGGTTAGTGATCGACAAACTAGTTTGTCTGTGACAAACTGTGATCCACA 2118
181 AATCAATTGAGATAGTTGGAATCTGAATCTCGAAGGGAATGGAGTGGCAACTGAACGTGCCA 240
2119 AATCAATTGAGATAGTTGGAATCTGAATCTCGAAGGGAATGGAGTGGCAACTGAACGTGCCA 2178
241 TCTGCAATTAAGATGGGCTTCAAGTCCGGTGTCCACCAAAAGGTGGTCAATTAATGAA 300
2179 TCTGCAATTAAGATGGGCTTCAAGTCCGGTGTCCACCAAAAGGTGGTCAATTAATGAA 2238
301 GCTGTGTAATGGCTGAAAACTGCTACAACTCTTGAAATCAAAAAACCTGACGGGAGTGAG 360
2239 GCTGTGTAATGGCTGAAAACTGCTACAACTCTTGAAATCAAAAAACCTGACGGGAGTGAG 2298
361 TGTCTACAGCAGCGCCAGACGGGAATTCGGGGCTTCCCGGGTCCCGGTATGTGCACAAA 420
2299 TGTCTACAGCAGCGCCAGACGGGAATTCGGGGCTTCCCGGGTCCCGGTATGTGCACAAA 2358
421 GTATCAGGAACGGGACCGTGTCCGGAGACTTTGCTTCCATAGAGGGTCTTTCTTC 480
2359 GTATCAGGAACGGGACCGTGTCCGGAGACTTTGCTTCCATAGAGGGTCTTTCTTC 2418
481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGTGAAGGTGTC 540
2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGTGAAGGTGTC 2478
541 GTTGCAATTTGATTAATGCCCCCAAGCTAAGAAAGGACTTCTTCAAGTCAACCCCTTGAGA 600
2479 GTTGCAATTTGATTAATGCCCCCAAGCTAAGAAAGGACTTCTTCAAGTCAACCCCTTGAGA 2538
601 GAGCCGGTCAATGCAACGAGGACCGTCTAGTGGCTACTATTCTACCAATTAAGATAT 660
2539 GAGCCGGTCAATGCAACGAGGACCGTCTAGTGGCTACTATTCTACCAATTAAGATAT 2598
661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
2599 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 2658
721 TAGCTCCAACTTGAATCAAGATTCAACACAGTTTCTGCTCCAGTGAATGAGACAATA 780
2659 TAGCTCCAACTTGAATCAAGATTCAACACAGTTTCTGCTCCAGTGAATGAGACAATA 2718
781 TATACAGTGGGAAGGAGCAATACACGGGAACCAATTAATTTGGAAGTCAACCCCGAA 840
2719 TATACAGTGGGAAGGAGCAATACACGGGAACCAATTAATTTGGAAGTCAACCCCGAA 2778
841 ATTGATACAACCAATCGGGAGTGGGCTTCTGGGAAACTAAAAAAACCTCACTAGAAAA 900
2779 ATTGATACAACCAATCGGGAGTGGGCTTCTGGGAAACTAAAAAAACCTCACTAGAAAA 2838
901 ATTGCAAGTGAAGAGTTGTCTTTCAAGTTGTATCAAAACGGAGCCAAAAACATCAAGTGT 960
2839 ATTGCAAGTGAAGAGTTGTCTTTCAAGTTGTATCAAAACGGAGCCAAAAACATCAAGTGT 2898
961 CAGAGTCCCGCGGAATCTTTCGACCCAGGACCAACACACACTGAAGACACACAA 1020
2899 CAGAGTCCCGCGGAATCTTTCGACCCAGGACCAACACACACTGAAGACACACAA 2958
1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1080
2959 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 3018
1081 GCAGTGTCCGATCTAAACACCCCTTGGCCAAATCTCCACGAGTCCCAATCTCACAACC 1140
3019 GCAGTGTCCGATCTAAACACCCCTTGGCCAAATCTCCACGAGTCCCAATCTCACAACC 3078
1141 AAACGAGTCCCGGACACGACGCCATATACACCCGTGTATAAATCTTGACATCTCTGAG 1200
3079 AAACGAGTCCCGGACACGACGCCATATACACCCGTGTATAAATCTTGACATCTCTGAG 3138

QY 1201 GCAACTCAAGTTGAACAACTATCACCAGAGAAACAGAAACGACAGACCTCCGACACT 1260
DB 3139 GCAACTCAAGTTGAACAACTATCACCAGAGAAACAGAAACGACAGACCTCCGACACT 3198
QY 1261 CCCTCTGCCACGACGCGGACCGGACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC 1320
DB 3199 CCCTCTGCCACGACGCGGACCGGACCCCAAAAGCAGAGAACCAACACGAGCAAGAGC 3258
QY 1321 ACTGACTTCTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAACAGCAGACCGCTGGC 1380
DB 3259 ACTGACTTCTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAACAGCAGACCGCTGGC 3318
QY 1381 AACAAACACTCATATCAACAAAGATACCGGAGAAAGAGTGCACGAGCGGGAAGTATGAGC 1440
DB 3319 AACAAACACTCATATCAACAAAGATACCGGAGAAAGAGTGCACGAGCGGGAAGTATGAGC 3378
QY 1441 TTAATTACCAATACTATTGCTGGAGTGCAGAGTGCATCAGAGCGGAGAGAACTCGA 1500
DB 3379 TTAATTACCAATACTATTGCTGGAGTGCAGAGTGCATCAGAGCGGAGAGAACTCGA 3438
QY 1501 CGA 1503
DB 3439 AGA 3441
RESULT 14
ACC71549
ID ACC71549 standard; DNA; 8199 BP.
XX ACC71549;
AC ACC71549;
XX 10-JUL-2003 (first entry)
XX VRC6604 (pAdApt Ebola GP(Z) (dTM)) plasmid.
DE Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Laesa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
OS Synthetic.
XX WO2003028632-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030251.
XX 01-OCT-2001; 2001US-0326476P.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Nabel GJ, Yang Z, Sullivan N, Sanchez A;
PI WPI; 2003-371961/35.
DR New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX Claim 1; Page 182-185; 219pp; English.
XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected

CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. The present sequence is an adenoviral shuttle vector
CC expressing Ebola virus glycoprotein (subtype Zaire) without its
CC transmembrane and intracellular domains
XX
SQ

Sequence 8199 BP; 1977 A; 2047 C; 2173 G; 2002 T; 0 U; 0 Other;

Query Match 66.7%; Score 1501.4; DB 10; Length 8199;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	ATGGCGGTTACAGGAAATATTCGAGTTACCTCGTGAATCAAGAGGACATCATCTTT	60
Db	1336	ATGGCGGTTACAGGAAATATTCGAGTTACCTCGTGAATCAAGAGGACATCATCTTT	1395
Qy	61	CTTTGGGTAATTTATCTCTTTTCCAAAGAACATTTTCATCCACCTTGGAGTCAACCAAT	120
Db	1396	CTTTGGGTAATTTATCTCTTTTCCAAAGAACATTTTCATCCACCTTGGAGTCAACCAAT	1455
Qy	121	AGCACATTACAGGTTAGTGTGACAAACTAGTTTGTCTGACAAACTGTCTATCCACA	180
Db	1456	AGCACATTACAGGTTAGTGTGACAAACTAGTTTGTCTGACAAACTGTCTATCCACA	1515
Qy	191	AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA	240
Db	1516	AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA	1575
Qy	241	TCTGCAACTAAAGATGGGCTTCAGGTCGGGTGCCACCAAAAGGTGGTCAATTTATGAA	300
Db	1576	TCTGCAACTAAAGATGGGCTTCAGGTCGGGTGCCACCAAAAGGTGGTCAATTTATGAA	1635
Qy	301	CTGTGTGAATGGCTGAAAACCTGCTACAACTTGAATCAAAAACCTGACGGAGTGAG	360
Db	1636	GCTGTGAATGGCTGAAAACCTGCTACAACTTGAATCAAAAACCTGACGGAGTGAG	1695
Qy	361	TGTTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGACAAA	420
Db	1696	TGTTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGACAAA	1755
Qy	421	GTATCAGGAACGGGACCGGTGTCGGGACATTTGCTTCCATTAAGAGGGTCTTCTTC	480
Db	1756	GTATCAGGAACGGGACCGGTGTCGGGACATTTGCTTCCATTAAGAGGGTCTTCTTC	1815
Qy	481	CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGC	540
Db	1816	CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGC	1875
Qy	541	GTTGCAATTTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA	600
Db	1876	GTTGCAATTTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA	1935
Qy	601	GAGCCGGTCAATGCAACCGAGGACCGGTCTAGTGGCTACTATTCTACCAATAGATAT	660
Db	1936	GAGCCGGTCAATGCAACCGAGGACCGGTCTAGTGGCTACTATTCTACCAATAGATAT	1995
Qy	661	CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC	720
Db	1996	CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC	2055
Qy	721	TACGTTCCAACTTGAATCAAGATTCTACCAACAGTTTCTGCTCCAGCTGAATGAGACAATA	780
Db	2056	TACGTTCCAACTTGAATCAAGATTCTACCAACAGTTTCTGCTCCAGCTGAATGAGACAATA	2115
Qy	781	TATACAAGTGGAAAAGAGGCAATACGAGGGAATACTAATTTGGAAGGTCAACCCCGAA	840
Db	2116	TATACAAGTGGAAAAGAGGCAATACGAGGGAATACTAATTTGGAAGGTCAACCCCGAA	2175
Qy	841	ATTGATACAACTCGGGAGTGGCCCTTCTGGGAAACTTAAAAAAAACTCACTAGAAAA	900
Db	2176	ATTGATACAACTCGGGAGTGGCCCTTCTGGGAAACTTAAAAAAAACTCACTAGAAAA	2235

Qy	901	ATTGCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAAAACGGAGCCAAAAACATCAGTGGT	960
Db	2236	ATTGCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAAAACGGAGCCAAAAACATCAGTGGT	2295
Qy	961	CAGAGTCCGGCGGAACCTTCTTCGAGCCAGGAGCAACAACTGAAGACCAAAA	1020
Db	2296	CAGAGTCCGGCGGAACCTTCTTCGAGCCAGGAGCAACAACTGAAGACCAAAA	2355
Qy	1021	ATCATGCTTCAGAAAAATTCCTTCGAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT	1080
Db	2356	ATCATGCTTCAGAAAAATTCCTTCGAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT	2415
Qy	1081	GCAGTGTGCATCTAAACCCCTTGCACCAATCTCCAGAGTCCCAATCCCTCACAACC	1140
Db	2416	GCAGTGTGCATCTAAACCCCTTGCACCAATCTCCAGAGTCCCAATCCCTCACAACC	2475
Qy	1141	AAACAGGTCGGGACAAACAGCACCAATAATACACCGTGTATAAATCTTGACATCTCTGAG	1200
Db	2476	AAACAGGTCGGGACAAACAGCACCAATAATACACCGTGTATAAATCTTGACATCTCTGAG	2535
Qy	1201	GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAAACAGCAGCAGCCTCCGACACT	1260
Db	2536	GCAACTCAAGTTGAACAACATCACCGCAGAACAGACAAACAGCAGCAGCCTCCGACACT	2595
Qy	1261	CCCTCTGCCAGCAGCGGACCGGACCCGCAAAAGCAGAGAAACCAACAGCAGCAGAGC	1320
Db	2596	CCCTCTGCCAGCAGCGGACCGGACCCGCAAAAGCAGAGAAACCAACAGCAGCAGAGC	2655
Qy	1321	ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAACAGCAGCAGCCTCGGC	1380
Db	2656	ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAAAACCAACAGCAGCAGCCTCGGC	2715
Qy	1381	AACAACAACTCATACCAAGATACCGGAGAAAGAGTGCCAGCAGCGGGAAGCTAGGC	1440
Db	2716	AACAACAACTCATACCAAGATACCGGAGAAAGAGTGCCAGCAGCGGGAAGCTAGGC	2775
Qy	1441	TTAATTACCAATACATTTGCTGGAGTGCAGGACTGATCAAGCGGGAAGAAAGACTCGA	1500
Db	2776	TTAATTACCAATACATTTGCTGGAGTGCAGGACTGATCAAGCGGGAAGAAAGACTCGA	2835
Qy	1501	CGA 1503	
Db	2836	AGA 2838	
RESULT 15			
ACC71548			
ID	ACC71548	standard; DNA; 8439 BP.	
XX	ACC71548;		
AC	ACC71548;		
XX	10-JUL-2003	(first entry)	
DT	10-JUL-2003		
XX	VRC6603	(pAdapt Ebola GP(Z)) plasmid.	
DE	Virucide; vaccine; immune response; Ebola virus; Marburg virus;		
XX	Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;		
KW	viral infection; filovirus; circular; cyclic; ds.		
XX	Synthetic.		
OS	WO2003028632-A2.		
XX	10-APR-2003.		
PN	24-SEP-2002; 2002WO-US030251.		
XX	01-OCT-2001; 2001US-0326476P.		
PD	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
PF	Nabel GJ, Yang Z, Sullivan N, Sanchez A;		
XX			

DR WPI; 2003-371961/35.
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX Claim 1; Page 180-182; 219pp; English.
PS
XX The present invention relates to a bimodal priming composition and
CC boosting competition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection
XX
SQ Sequence 8439 BP; 2059 A; 2081 C; 2210 G; 2089 T; 0 U; 0 Other;
Query Match 66.7%; Score 1501.4; DB 10; Length 8439;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGGCGTTACAGGAATATTGCGATTACCTCGTGATTCGATTCAAGAGGACATCAATCTTT 60
DB 1336 ATGGGCGTTACAGGAATATTGCGATTACCTCGTGATTCGATTCAAGAGGACATCAATCTTT 1395
QY 61 CTTTGGGTAATATTCCTTTTCCAAAGAACATTTTCCATCCCATCTTGGAGTCATCCCAAT 120
DB 1396 CTTTGGGTAATATTCCTTTTCCAAAGAACATTTTCCATCCCATCTTGGAGTCATCCCAAT 1455
QY 121 ASCACATTAACAGTTAGTATGTCGACAACTAGTTTGTGTCGACAACTGTCATCCCA 180
DB 1456 AGCACATTAACAGTTAGTATGTCGACAACTAGTTTGTGTCGACAACTGTCATCCCA 1515
QY 181 AATCAATTGAGATCAGTTGGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 240
DB 1516 AATCAATTGAGATCAGTTGGAATCTCGAAGGGAATGGAGTGGCAACTGACGTGCCA 1575
QY 241 TCTGCAACTAAAGATGGGGCTTCAGGTCGGTGTCCCAACAAAGGTGTCAATATGAA 300
DB 1576 TCTGCAACTAAAGATGGGGCTTCAGGTCGGTGTCCCAACAAAGGTGTCAATATGAA 1635
QY 301 GCTGGTGAATGGGCTGAAACTGCTACAACTTTGAAATCAAAAACCTGACGGGAGTGAG 360
DB 1636 GCTGGTGAATGGGCTGAAACTGCTACAACTTTGAAATCAAAAACCTGACGGGAGTGAG 1695
QY 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGGTGGCGGTATGTGCACAAA 420
DB 1696 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGGTGGCGGTATGTGCACAAA 1755
QY 421 GTATCAGGAACGGGAACCGTGTGCGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 480
DB 1756 GTATCAGGAACGGGAACCGTGTGCGGAGACTTTTGCCTTCCATAAAGAGGGTCTTTCTTC 1815
QY 481 CTGTATGATCGACTTCTTCCACAGTTATCTACCGAGGAACGACTTTCCTCAAGGTGTC 540
DB 1816 CTGTATGATCGACTTCTTCCACAGTTATCTACCGAGGAACGACTTTCCTCAAGGTGTC 1875
QY 541 GTTGCAATTTCTGATACTGCCCCAAGCTAAGAAGGACTTTCTTCAAGTCAACCCCTTGGA 600
DB 1876 GTTGCAATTTCTGATACTGCCCCAAGCTAAGAAGGACTTTCTTCAAGTCAACCCCTTGGA 1935
QY 601 GAGCCGGTCAATGCAACGGAGACCGGCTCTAGTGGCTACTATTCTACCACAATTTAGAT 660
DB 1936 GAGCCGGTCAATGCAACGGAGACCGGCTCTAGTGGCTACTATTCTACCACAATTTAGAT 1995

QY 661 CAGGCTACCGCTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGTTTGACAAATTTGACC 720
DB 1996 CAGGCTACCGCTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGTTTGACAAATTTGACC 2055
QY 721 TACGTCCAACTTGAATCAAGATTACACCAACAGTTTCTGTCTCCAGCTGAATGAGACAATA 780
DB 2056 TACGTCCAACTTGAATCAAGATTACACCAACAGTTTCTGTCTCCAGCTGAATGAGACAATA 2115
QY 781 TATACAGTGGGAAAGAGGACCAATACACCGGAAACCTAATTTGGAAGGTCAACCCCGAA 840
DB 2116 TATACAGTGGGAAAGAGGACCAATACACCGGAAACCTAATTTGGAAGGTCAACCCCGAA 2175
QY 841 ATTGATACAACTCGGGGAGTGGGCTTCTTGGGAACTAATAAATAAATCTCCTACATAGAAA 900
DB 2176 ATTGATACAACTCGGGGAGTGGGCTTCTTGGGAACTAATAAATAAATCTCCTACATAGAAA 2235
QY 901 ATTGAGTGAAGAGTGTCTTTTCAAGTTGATCAAAACGGAGCCAAACATCATGTGGT 960
DB 2236 ATTGAGTGAAGAGTGTCTTTTCAAGTTGATCAAAACGGAGCCAAACATCATGTGGT 2295
QY 961 CAGAGTCCGCGCGCAACTTCTTCCGACCCAGGACCAACACAACTGAGACCCACAAA 1020
DB 2296 CAGAGTCCGCGCGCAACTTCTTCCGACCCAGGACCAACACAACTGAGACCCACAAA 2355
QY 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1080
DB 2356 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT 2415
QY 1081 GCAGTGTCCGATCTAAACAACTTCCGACCAATCTCCAGAGTCCCAATCCCTCAACACC 1140
DB 2416 GCAGTGTCCGATCTAAACAACTTCCGACCAATCTCCAGAGTCCCAATCCCTCAACACC 2475
QY 1141 AAAACAGTCCGCGACACAGCACCCATTAATACACCGTGTATTAACCTTGACATCTCTGAG 1200
DB 2476 AAAACAGTCCGCGACACAGCACCCATTAATACACCGTGTATTAACCTTGACATCTCTGAG 2535
QY 1201 GCNACTCAAGTTGAACAACTACCGCAGAAACAGACAAACGACAGCAGACCTCCGACACT 1260
DB 2536 GCNACTCAAGTTGAACAACTACCGCAGAAACAGACAAACGACAGCAGACCTCCGACACT 2595
QY 1261 CCCTTGCACGACGCGAGCCGACCCCAAAAGAGAGAAACCAACACGAGCAAGAGC 1320
DB 2596 CCCTTGCACGACGCGAGCCGACCCCAAAAGAGAGAAACCAACACGAGCAAGAGC 2655
QY 1321 ACTGACTTCTTGGACCCCGCCACACACAAAGTCCCAAAACCCAGAGAGACCCCTGGC 1380
DB 2656 ACTGACTTCTTGGACCCCGCCACCAACAAAGTCCCAAAACCCAGAGAGACCCCTGGC 2715
QY 1381 AACAAACACCTCATCACAAAGATACCGGAGAAAGAGTGCAGCAGCGGGAAGCTTAGGC 1440
DB 2716 AACAAACACCTCATCACAAAGATACCGGAGAAAGAGTGCAGCAGCGGGAAGCTTAGGC 2775
QY 1441 TTAATTAACAATACTATTGCTGGAGTCCAGAGCTGATCAAGGCGGAGAGAACTCGA 1500
DB 2776 TTAATTAACAATACTATTGCTGGAGTCCAGAGCTGATCAAGGCGGAGAGAACTCGA 2835
QY 1501 CGA 1503
DB 2836 AGA 2838

Search completed: July 31, 2005, 16:57:12
Job time : 1185.79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:13:05 ; Search time 7351.67 Seconds
(without alignments)
11660.032 Million cell updates/sec

Title: US-10-066-506A-1
Perfect score: 2252
Sequence: 1 atggcgcttacaggaattt.....tatatcgataacggaattc 2252

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
c 1	77	3.4	895	9	CNS0071A	AL066286	Drosophila
c 2	70.4	3.1	922	9	CNS0073W	AL066784	Drosophila
c 3	67.8	3.0	884	9	CNS006U0	AL065923	Drosophila
c 4	66.6	3.0	909	9	CNS00JTL	AL076720	Drosophila
c 5	62.6	2.8	1101	9	CNS00FXE	AL071370	Drosophila
c 6	60.4	2.7	902	9	CNS00QOP	AL065804	Drosophila
c 7	58.4	2.6	1101	9	CNS01J2T	AL080707	Drosophila
c 8	58.4	2.6	1225	9	CNS0166K	AL106358	Drosophila
c 9	57.2	2.5	918	9	AG405745	AG405745	Mus muscu
c 10	55.4	2.5	937	9	CNS006ST	AL065880	Drosophila
c 11	55.2	2.5	962	9	AG126141	AG126141	Pan trogl
c 12	54.4	2.4	939	9	CNS00CNG	AL059400	Drosophila
c 13	53.4	2.4	513	9	CNS02CMF	AL191328	Tetraodon
c 14	53.2	2.4	1806	9	CG756794	CG756794	P051-4-G0
c 15	53	2.4	934	9	CNS00B4F	AL060213	Drosophila
c 16	53	2.4	1844	9	CG756580	CG756580	P051-4-B0
c 17	52.6	2.3	1051	9	CNS015H2	AL105440	Drosophila
c 18	52.6	2.3	1101	9	CNS00LT2	AL078714	Drosophila
c 19	52.4	2.3	718	9	AG137045	AG137045	Pan trogl
c 20	52.2	2.3	788	9	AG125754	AG125754	Pan trogl
c 21	52.2	2.3	800	9	AG378374	AG378374	Mus muscu
c 22	52	2.3	733	9	CNS011AC	AL100014	Drosophila
c 23	51.6	2.3	626	6	BY752508	BY752508	
c 24	51.2	2.3	527	9	CE480293	CE480293	tigr-gss-

25	50.8	2.3	1101	9	CNS00LO0	AL068607	Drosophila
c 26	50.8	2.3	1101	9	CNS017RJ	AL108409	Drosophila
c 27	50.6	2.2	969	6	CD325655	CD325655	AGENCOURT
c 28	50.6	2.2	1101	9	CNS00Z1D	AL097099	Drosophila
c 29	50.4	2.2	1256	5	BQ433505	BQ433505	AGENCOURT
c 30	50.2	2.2	868	5	BU797633	BU797633	SJF2EPC03
c 31	49.6	2.2	993	9	AG136567	AG136567	Pan trogl
c 32	49.6	2.2	1101	9	CNS017JU	AL108132	Drosophila
c 33	49.6	2.2	1752	9	CG756831	CG756831	P051-4-H0
c 34	49.4	2.2	572	9	CG897362	CG897362	pastbac04
c 35	49.4	2.2	790	9	CNS0379X	AL231054	Tetraodon
c 36	49.2	2.2	1022	8	BZ561449	BZ561449	pacs2-164
c 37	49.2	2.2	1087	9	CNS0106Y	AL098596	Drosophila
c 38	49	2.2	645	9	AG136205	AG136205	Pan trogl
c 39	48.8	2.2	1101	9	CNS017WI	AL108588	Drosophila
c 40	48.6	2.2	804	4	B1948870	B1948870	HVSME1001
c 41	48.6	2.2	926	8	AQ782232	AQ782232	HS_3176_B
c 42	48.4	2.1	1103	9	AG135574	AG135574	Pan trogl
c 43	48.4	2.1	1155	9	AG030839	AG030839	Pan trogl
c 44	48.2	2.1	712	9	CNS04AL2	AL281999	Tetraodon
c 45	48	2.1	714	9	CL855158	CL855158	OR_Cha008

ALIGNMENTS

RESULT 1
CNS0071A/c
LOCUS
DEFINITION
CNS0071A 895 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
AL066286
AL066286.1 GI:4945153
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1 (bases 1 to 895)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org/TheBDGP/Drosophila> melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosier in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..895
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14B09"
/note="end : TET3"

ORIGIN

Query Match 3.4%; Score 77; DB 9; Length 895;
Best Local Similarity 22.4%; Pred. No. 1.2e-10;

[illegible]

RESULT 2	CNS0073W/c	922 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	BACR14D09	of RPC1-98	library	from <i>Drosophila melanogaster</i>	(fruit fly), genomic survey sequence.
DEFINITION	AL066784.1	GI:4945247	GSS.		
ACCESSION	AL066784.1	GI:4945247	GSS.		
VERSION	AL066784.1	GI:4945247	GSS.		
KEYWORDS					
SOURCE	<i>Drosophila melanogaster</i>	(fruit fly)			
ORGANISM	<i>Drosophila melanogaster</i>				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; <i>Drosophila</i> .				
REFERENCE	1	(bases 1 to 922)			
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr) - Web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley <i>Drosophila</i> Genome Project (BDGP). The BDGP is constructing a physical map of the <i>Drosophila melanogaster</i> genome using these BACs. For further information please see http://www.fruitfly.org The BDGP <i>Drosophila melanogaster</i> BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPC1-98 and was constructed by partial EcoRI digestion of <i>Drosophila</i> DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm . Location/Qualifiers				
FEATURES	1..922				
source	/organisms="Drosophila melanogaster"				

[illegible]

RESULT 3	CNS006U0	884 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence T7 end of BAC #				
DEFINITION	BACR14N21 of RCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCSSION	AL065923				
VERSION	AL065923.1	GI:4944891			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 884)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.hcm .				
FEATURES	Location/Qualifiers				
	1..884				

REFERENCE AUTHORS TITLE JOURNAL	COMMENT
Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1101) Genoscope. Direct Submission Submitted (23-JUL-1999) BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - web : www.genoscope.cns.fr)	Genoscope - Centre National de Sequencage : Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaut at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBel0BAC11.

FEATURES	source	Location/Qualifiers
pBelobAC11	1. .il01	/organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACN37N10" /clone_lib="DrosBAC" /plasmid="pBelobAC11" /note="end : T7"

ORIGIN

Query Match	2.6%;	Score 58.4;	DB 9;	Length 1101;
Best Local Similarity	36.4%;	Pred. No. 4e-05;		
Matches 102;	Conservative 63;	Mismatches 109;	Indels 6;	Gaps 1;
Qy	1124	CCCAATCCCTCACAAACCAACCAGGTCGGAGACAACAGCACCACCCATAATACACCCGTGTATA	1183	
Db	1099	CMCCCCMCMCMCCAMCMCAACCCCCMCMCAAMMCAACCCCMMAAMMAAMMM-----A	1046	
Qy	1184	AACTTGACATCTCTGTGGGCAACTCAAGTTGAAACAACATCATCCGCAGACAACAGACAACGACA	1243	
Db	1045	CAMCCCAACMCCAMANAACAMMAAMMAACMAAAACMAAAACMAACMAAAMCMAMAMCM	986	
Qy	1244	GCACAGCCTCCGACACTCCCTCTGTCCACAGCAGCCGCGGACCCCAAAAGCAGAGAACA	1303	
Db	985	MCAMAMAMMCMMAACMCMCAACCCCMCMCMCMAMMCMCAACMAAMCMAMAMCM	926	
Qy	1304	CCAAACAGGACAAGACGACTGCTTCTGGACCCCGCCACCAACAACGATCCCCCAAAACC	1363	
Db	925	MCAACAMCYCAMCCAAAMACAACAMACACCAACCCCCCMCMMAAAACCMMAAAAAAMMM	866	
Qy	1364	ACAGCGAGACCGTGGCAACAACAACACTCATCAACCAAGA	1403	
Db	865	AACHCCCAACWCAACMAAAAAACAAAAACAAAAAAMA	826	

RESULT 8

CNS0166K/c	CNS0166K	1225 bp	DNA	linear	GSS 26-JUL-1999
LOCUS					
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC BACN15C20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				

ACCESSION
AL106358

VERSION AL106358.1 GI:5621512

KEYWORDS: GSS.

SOURCE *Drosophila melanogaster* (fruit fly)

ORGANISM *Drosophila melanogaster*

Eukaryota; Metazoa; Arthropoda; Hex-

Neoptera; Endopterygota; Diptera; E

Ephydroidea; Drosophilidae; Drosophila

REFERENCE 1 (bases 1 to 1225)

AUTHORS Genoscope, . . .

TITLE	Direct Submission
1. TITLE	1. TITLE
2. AUTHOR	2. AUTHOR
3. ABSTRACT	3. ABSTRACT
4. KEYWORDS	4. KEYWORDS
5. REFERENCES	5. REFERENCES
6. COMMENTS	6. COMMENTS
7. INDEXING	7. INDEXING
8. SUMMARY	8. SUMMARY
9. CONCLUSIONS	9. CONCLUSIONS
10. REFERENCES	10. REFERENCES
11. COMMENTS	11. COMMENTS
12. INDEXING	12. INDEXING
13. SUMMARY	13. SUMMARY
14. CONCLUSIONS	14. CONCLUSIONS
15. REFERENCES	15. REFERENCES
16. COMMENTS	16. COMMENTS
17. INDEXING	17. INDEXING
18. SUMMARY	18. SUMMARY
19. CONCLUSIONS	19. CONCLUSIONS
20. REFERENCES	20. REFERENCES
21. COMMENTS	21. COMMENTS
22. INDEXING	22. INDEXING
23. SUMMARY	23. SUMMARY
24. CONCLUSIONS	24. CONCLUSIONS
25. REFERENCES	25. REFERENCES
26. COMMENTS	26. COMMENTS
27. INDEXING	27. INDEXING
28. SUMMARY	28. SUMMARY
29. CONCLUSIONS	29. CONCLUSIONS
30. REFERENCES	30. REFERENCES
31. COMMENTS	31. COMMENTS
32. INDEXING	32. INDEXING
33. SUMMARY	33. SUMMARY
34. CONCLUSIONS	34. CONCLUSIONS
35. REFERENCES	35. REFERENCES
36. COMMENTS	36. COMMENTS
37. INDEXING	37. INDEXING
38. SUMMARY	38. SUMMARY
39. CONCLUSIONS	39. CONCLUSIONS
40. REFERENCES	40. REFERENCES
41. COMMENTS	41. COMMENTS
42. INDEXING	42. INDEXING
43. SUMMARY	43. SUMMARY
44. CONCLUSIONS	44. CONCLUSIONS
45. REFERENCES	45. REFERENCES
46. COMMENTS	46. COMMENTS
47. INDEXING	47. INDEXING
48. SUMMARY	48. SUMMARY
49. CONCLUSIONS	49. CONCLUSIONS
50. REFERENCES	50. REFERENCES
51. COMMENTS	51. COMMENTS
52. INDEXING	52. INDEXING
53. SUMMARY	53. SUMMARY
54. CONCLUSIONS	54. CONCLUSIONS
55. REFERENCES	55. REFERENCES
56. COMMENTS	56. COMMENTS
57. INDEXING	57. INDEXING
58. SUMMARY	58. SUMMARY
59. CONCLUSIONS	59. CONCLUSIONS
60. REFERENCES	60. REFERENCES
61. COMMENTS	61. COMMENTS
62. INDEXING	62. INDEXING
63. SUMMARY	63. SUMMARY
64. CONCLUSIONS	64. CONCLUSIONS
65. REFERENCES	65. REFERENCES
66. COMMENTS	66. COMMENTS
67. INDEXING	67. INDEXING
68. SUMMARY	68. SUMMARY
69. CONCLUSIONS	69. CONCLUSIONS
70. REFERENCES	70. REFERENCES
71. COMMENTS	71. COMMENTS
72. INDEXING	72. INDEXING
73. SUMMARY	73. SUMMARY
74. CONCLUSIONS	74. CONCLUSIONS
75. REFERENCES	75. REFERENCES
76. COMMENTS	76. COMMENTS
77. INDEXING	77. INDEXING
78. SUMMARY	78. SUMMARY
79. CONCLUSIONS	79. CONCLUSIONS
80. REFERENCES	80. REFERENCES
81. COMMENTS	81. COMMENTS
82. INDEXING	82. INDEXING
83. SUMMARY	83. SUMMARY
84. CONCLUSIONS	84. CONCLUSIONS
85. REFERENCES	85. REFERENCES
86. COMMENTS	86. COMMENTS
87. INDEXING	87. INDEXING
88. SUMMARY	88. SUMMARY
89. CONCLUSIONS	89. CONCLUSIONS
90. REFERENCES	90. REFERENCES
91. COMMENTS	91. COMMENTS
92. INDEXING	92. INDEXING
93. SUMMARY	93. SUMMARY
94. CONCLUSIONS	94. CONCLUSIONS
95. REFERENCES	95. REFERENCES
96. COMMENTS	96. COMMENTS
97. INDEXING	97. INDEXING
98. SUMMARY	98. SUMMARY
99. CONCLUSIONS	99. CONCLUSIONS
100. REFERENCES	100. REFERENCES
101. COMMENTS	101. COMMENTS
102. INDEXING	102. INDEXING
103. SUMMARY	103. SUMMARY
104. CONCLUSIONS	104. CONCLUSIONS
105. REFERENCES	105. REFERENCES
106. COMMENTS	106. COMMENTS
107. INDEXING	107. INDEXING
108. SUMMARY	108. SUMMARY
109. CONCLUSIONS	109. CONCLUSIONS
110. REFERENCES	110. REFERENCES
111. COMMENTS	111. COMMENTS
112. INDEXING	112. INDEXING
113. SUMMARY	113. SUMMARY
114. CONCLUSIONS	114. CONCLUSIONS
115. REFERENCES	115. REFERENCES
116. COMMENTS	116. COMMENTS
117. INDEXING	117. INDEXING
118. SUMMARY	118. SUMMARY
119. CONCLUSIONS	119. CONCLUSIONS
120. REFERENCES	120. REFERENCES
121. COMMENTS	121. COMMENTS
122. INDEXING	122. INDEXING
123. SUMMARY	123. SUMMARY
124. CONCLUSIONS	124. CONCLUSIONS
125. REFERENCES	125. REFERENCES
126. COMMENTS	126. COMMENTS
127. INDEXING	127. INDEXING
128. SUMMARY	128. SUMMARY
129. CONCLUSIONS	129. CONCLUSIONS
130. REFERENCES	130. REFERENCES
131. COMMENTS	131. COMMENTS
132. INDEXING	132. INDEXING
133. SUMMARY	133. SUMMARY
134. CONCLUSIONS	134. CONCLUSIONS
135. REFERENCES	135. REFERENCES
136. COMMENTS	136. COMMENTS
137. INDEXING	137. INDEXING
138. SUMMARY	138. SUMMARY
139. CONCLUSIONS	139. CONCLUSIONS
140. REFERENCES	140. REFERENCES
141. COMMENTS	141. COMMENTS
142. INDEXING	14

JOURNAL
Submitted (23-JUL-1999) Genoscope -

BP 191 91006 EVRY cedex - FRANCE (E)

- Web : www.genoscope.cns.fr

COMMENT: Determination of this BAC-end sequence collaboration with the European Drug

collaboration with the European Drug

http://www.edgpb.ebi.ac.uk -. This *Drosophila melanogaster* BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.

```
FEATURES
    source
        1..1225
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACN15C20"
            /clone_lib="DrosBAC"
            /plasmid="pBelobAC11"
            /note="end : SP6"
```

ORIGIN

[illegible]

RESULT 9

AG405745	918 bp	DNA	linear	GSS 03-JUN-2000
Mus musculus molossinus	DNA, clone:MSMg01-264B16.TJ,	genomic survey		
sequence.				
AG405745				
AG405745.1	GI:48048431			
GSS.				
Mus musculus molossinus				
Mus musculus molossinus				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.				
1				
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.				
BAC end Sequences of Library MSMg01				
2				
(bases 1 to 918)				
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.				
Direct Submission				
1				
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.				
2				
(bases 1 to 918)				
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.				
Direct Submission				

JOURNAL Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: hattori@gsc.riken.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the mouse BAC library MSMG01. For BAC library availability, please contact Kuniya Abe (abe@rtc.riken.jp). Tsukuba Institute, Bio Resource Center, The Institute of Physical and Chemical Research (RIKEN) 3-1-1 Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY Vector : pBACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI
FEATURES Location/Qualifiers
source 1..918
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMG01-264B16.TJ"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_lib="MSMG01 Mouse Male BAC Library"
ORIGIN
Query Match 2.5%; Score 57.2; DB 9; Length 918;
Best Local Similarity 52.1%; Pred. No. 8.6e-05;
Matches 148; Conservative 0; Mismatches 135; Indels 1; Gaps 1;
QY 1124 CCCATCCCTCACAACCAACAGGTCCGGACACAGACCCCATATATACACCGGTGATAT 1183
DB 865 CCCAGACCAACGACCGCCCGCGGAGAAACACACACACACCCCGCCCGCCATTA 806
QY 1184 AACTTGACATCTCTGAGGCACTCAAGTTGAAACAACATCACCGCAGACAGACAACGACA 1243
DB 805 GACCGACCAACCCCAACACCCACAGAACACACACCATGACACACACAGATCCACA 746
QY 1244 GCACAGCTCCGACATCCCTCTGCGACGACCGGACCCCGCCAAAGCAGAGAACA 1303
DB 745 AGCAGCGCCCGACACAACCAACATAGCTCGACCCACAGACACACACCCCGCT 686
QY 1304 CCACACAGCAGACGACGACTGCTTCTGACCGCCCGCCACACACAGTCCCGCAACCC 1363
DB 685 CCAACCCCAAGCGGACCACTCACTACCCAGATACGCGCCACCTAACCCACCCCGCC 627
QY 1364 ACAGCGAGACCGCTGGCAACAACAACATCTCATCCACCAAGATACC 1407
DB 625 ACAGCCACACCACTCCCGCCACACACCAACCAACCAACCAACCAAC 583
RESULT 10
CNS006ST 937 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC # BAC14F16 of Rpci-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
LOCUS AL065880
ACCESSION AL065880.1 GI:4944848
VERSION
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscormorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 937)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named Rpci-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES Location/Qualifiers
source 1..937
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BAC14F16"
/clone_lib="Rpci-98"
/note="end : TET3"
ORIGIN
Query Match 2.5%; Score 55.4; DB 9; Length 937;
Best Local Similarity 21.6%; Pred. No. 0.00029;
Matches 88; Conservative 140; Mismatches 179; Indels 0; Gaps 0;
QY 1092 TCTAACACCCCTTGCCACAATCTCCACGAGTCCCCCAATCTCCACACCAACACAGGTCC 1151
DB 437 TTTTYYMCMACMAWMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAA 496
QY 1152 GGACACAGCAGCCCAATATACACCGTGTATAAATTGACATCTCTGAGGCACTCAAGT 1211
DB 497 MCMWACACCMWMMCMWMMCMWMMCMWMMCMWMMCMWMMCMWMMCMWMMCMWMMCMW 556
QY 1212 TGACAAACATCACCGCAACACAGACAGACAGCAGCAGCAGCAGCAGCAGCAGCAGC 1271
DB 557 CMWMAACMCAWMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMA 616
QY 1272 GACCGCAGCCGAGCCCCCAAGCAGAGAGAACACCAACAGCAGCAGCAGCAGCAGCAGC 1331
DB 617 AWAAMAAMMAWCMWCMWCMWCMWCMWCMWCMWCMWCMWCMWCMWCMWCMWCMWCMW 676
QY 1332 GGACCCCGCCACACACAGTCCCCCAACACACAGCAGCAGCAGCAGCAGCAGCAGCAG 1391
DB 677 MWCMAAAVMWGCACVCWMAARMVMWMMVMWMAAMVMWMAAMVMWMAAMVMWMAAMVM 736
QY 1392 TCATCACCAGATACCGGAGAGAGAGTGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1451
DB 737 AMWCMRAAASAGASASRRRRGAACVVRGVGSARGSGVVGMRSGAGMARCAVMG 796
QY 1452 TACTATTGCTGGAGTCGAGGACTGATCAGACGGCGGAGAGAACTC 1498
DB 797 VAVRVASRAMRAASAGARVWMSGVMSACRCVWMSMAAAAHAMTATM 843
RESULT 11
AG126141 962 bp DNA linear GSS 04-NOV-2001
LOCUS
DEFINITION Pan troglodytes DNA, clone: PTB-136J19.F, genomic survey sequence.
ACCESSION AG126141.1 GI:16655306
VERSION
KEYWORDS
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki Y., Watanabe H. and Sakaki Y.
BAC end sequences of Library PTB
2 (bases 1 to 962)
Fujiyama A., Hattori M., Toyoda A., Taylor T.D., Yada T.,
Totoki Y., Watanabe H. and Sakaki Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbesgsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. .962
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-136519.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

Query Match 2.5%; Score 55.2; DB 9; Length 962;
Best Local Similarity 55.1%; Pred. No. 0.00034;
Matches 108; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1204 ACTCAAGTTGACACATCACCAGGAGACAGACAGACAGACGCTCCGACACTCC 1263
Db ACCCAAAAAGACCCCAACACGACCAACCAACCAACCAACCAACCAACCAACCA 751

QY 1264 TCTGCCAGCAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAG 1323
Db CAATCTCAGACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 811

QY 1324 GACTTCTGGACCCCGCCACCAACAGTCCCAACCAACCAACCAACCAACCGCTGGCAAC 1383
Db AACCCTCCCAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 871

QY 1384 ACAACACTCATCACC 1399
Db AAAAAAACAACCAACCC 887

RESULT 12
CNS00CNG 939 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BAC26H16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL059400
VERSION AL059400.1 GI:4946964
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 939)
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1. .939
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR26H16"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 2.4%; Score 54.4; DB 9; Length 939;
Best Local Similarity 12.8%; Pred. No. 0.00058;
Matches 36; Conservative 143; Mismatches 103; Indels 0; Gaps 0;

QY 1093 CTAACAACCTTGGCACAATCTCCAGAGTCCCAATCCCTCACAACCAACCAACGAGTCCG 1152
Db CTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 346

QY 1153 GACAACAGCACCATAATACACCGGTGTATAACTTGACATCTCTGAGGCAACTCAAGTT 1212
Db CCYATCCCTCMYHTTYMCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 406

QY 1213 GAACAACATACCGCAGACAGACAGACAGACAGACAGACAGACAGACAGACAGTCCCTCTGCCACG 1272
Db MMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMM 466

QY 1273 ACCGACGCGGACCCCAACAGAGAGACCAACCAACCAACCAACGACGACGACTCTCTCG 1332
Db MMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMM 526

QY 1333 GACCCGCGCACCAACAAGTCCCAACCAACCAACCAACGAGAGAC 1374
Db MMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMM 568

RESULT 13
CNS02CMF 513 bp DNA linear GSS 01-SEP-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 255P15 of library G from tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL191328
VERSION AL191328.1 GI:7829432
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodontidae.
1
Roest Croliius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, P.,
Saurin, W. and Weissenbach, J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
20296633

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE


```

10835645
2 Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
REFERENCE Fizames,C., Flecher,C., Bouneau,L., Billault,A., Quetier,F.,
AUTHORS Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)
MEDLINE
PUBMED 20359837
REFERENCE 10899143
AUTHORS 3 (bases 1 to 513)
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
1..513
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone_lib="255P15"
/clone_lib="G"
/note="Genoscope sequence ID : C0AG255CH08LPI-end : T7"
ORIGIN
Query Match 2.4%; Score 53.4; DB 9; Length 513;
Best Local Similarity 27.0%; Pred. No. 0.00094;
Matches 84; Conservative 88; Mismatches 139; Indels 0; Gaps 0;
QY 1100 CCGTTGCCAATCTCCAGAGTCCCAATCCCTCACAACCAACAGCGTCCGGACAACA 1159
DB 444 CCMTTACTTTTTTTTAAACTTCCNAACATCMAMCMCCAMWATWACAMCMWAMA 385
QY 1160 GCACCATATACACCCGTTGTAATACTTGACATCTCTGAGGCAACTCAAGTTGAACAAC 1219
DB 384 AANTMMCMWAGAMCMCAAAAAMAMCMCMACCMWATWAAAMMMWAGACMACWAS 325
QY 1220 ATCACCGCAGACAGACAAACGACAGACGCTCCGACACTCCCTCTGCGACGCGCAG 1279
DB 324 ACCMCAWAGAMCMCMWATACCMWATMACAMWAMWAGACCMWACMACACCMWATAA 265
QY 1280 CCGGACCCCAAGAGCAGACAGACCAACAGCAGCAGAGAGAGTACTGACTCTCTGGACCCCG 1339
DB 264 AMACMCMTWAGAMCMCMWATACCMWATMACAMWAMWAGACCMWAMWAMWAMWMMMG 205
QY 1340 CCACCAACAAGTCCCAAAACCAACAGCAGACGCTGGCAACACACACTCATCACC 1399
DB 204 AMAMMMWAGACCMWATMACCMWACMASAMMMWAMWAMWAMWAMWAMWAMWAMWAMW 145
QY 1400 AAGATACCGGA 1410
DB 144 ACACCAACCA 134
RESULT 14
CG756794/c
LOCUS 1806 bp DNA linear GSS 24-OCT-2003
DEFINITION P051-4-G04.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,
genomic survey sequence.
ACCESSION CG756794
VERSION CG756794.1 GI:37984708
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1806)
AUTHORS Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,

```

```

Buntjer,J., van der Meulen,M. and Sommer,R.J.
An integrated physical and genetic map of the nematode Pristionchus
pacificus
Mol. Genet. Genomics 269 (5), 715-722 (2003)
22835951
12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES
source
1..1806
Location/Qualifiers
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."
ORIGIN
Query Match 2.4%; Score 53.2; DB 9; Length 1806;
Best Local Similarity 53.3%; Pred. No. 0.0016;
Matches 153; Conservative 0; Mismatches 131; Indels 3; Gaps 2;
QY 1113 CTCACGAGTCCCAATCCCTCACAACCAACAGAGTCCGGACACAGACCCATAATAC 1172
DB 1610 CACCACAACACGGAACCAACAANCCACCACCAACAAACAAACACACACACCA 1551
QY 1173 ACCCGTGTATAACTTGATCTCTGAGGCAACTCAAGTTGAACAACATCACCGCAGAAC 1232
DB 1550 ACACCCACAACAACAACAAAGACAGACACACCCACACACACACACACACACAC 1491
QY 1233 AGACAACGACAGCAGCGCTCCGACACTCTCTGCGACGACGCGGCGGACCCCAAA 1292
DB 1490 AAACNAAACCAACACAC--ACCACACCCACCCACCCACACACACCGGCCACCC 1433
QY 1293 AGCAGAGAACCAAC--CAGCAGCAGAGACACTGACTTCTTGAGCCCGCCACCAACAA 1351
DB 1432 ACCCCACAACCAACCCACCAACACACCCACCGCCACCAACCCCGCCACCAACAC 1373
QY 1352 GTCCCCAAACCAACGACGCGCTGGCAACAACACACTCATCAC 1398
DB 1372 CCCCCCAACAGACCAACAAACCCCAAGCACCACCCCAACACAC 1326
RESULT 15
CNS000D4F 934 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR26H07 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL060213
VERSION AL060213.1 GI:4947565
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 934)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).

```


Search completed: August 1, 2005, 10:09:02
Job time : 7356.67 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:35:29 ; Search time 360.705 Seconds
(without alignments)
10215.808 Million cell updates/sec

Title: US-10-066-506A-1

Perfect score: 2252

Sequence: 1 atgggcgttacaggaattt.....tatatcgataacggaattc 2252

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1501.4	66.7	2172	3	US-08-760-615-1
2	1499.8	66.6	2298	4	US-09-650-086A-1
3	742.2	33.0	11460	4	US-09-336-910A-1
4	740.2	32.9	2164	3	US-08-760-615-3
5	545.2	24.2	2247	3	US-08-760-615-5
6	49.2	2.2	399	4	US-09-621-976-8976
7	46.8	2.1	194889	4	US-09-949-016-15654
8	46.4	2.1	34230	4	US-09-949-016-12052
9	46.4	2.1	128470	4	US-09-949-016-13765
10	44.8	2.0	152132	4	US-09-949-016-13845
11	44.8	2.0	152145	4	US-09-949-016-12371
12	44.4	2.0	7218	1	US-08-232-463-14
13	44.2	2.0	39456	4	US-09-949-016-14613
14	43.8	1.9	522	4	US-09-949-016-103758
15	43.8	1.9	39154	4	US-09-949-016-12384
16	43.8	1.9	39154	4	US-09-949-016-12801
17	43.8	1.9	39443	4	US-09-949-016-14326
18	43.8	1.9	39443	4	US-09-949-016-14327
19	42	1.9	4403765	3	US-09-103-840A-2
20	42	1.9	4411529	3	US-09-103-840A-1
21	41.4	1.8	1968	4	US-09-248-796A-1729
22	41	1.8	242	4	US-09-354-147C-30
23	40.2	1.8	83617	4	US-09-949-016-12254
24	39.6	1.8	271134	4	US-09-949-016-12705
25	39.6	1.8	305491	4	US-09-949-016-17550
26	38.6	1.7	570	4	US-09-248-796A-7987
27	38.6	1.7	67620	4	US-09-949-016-16939

28	38.6	1.7	114793	4	US-10-148-806-3	Sequence 3, Appli
29	38.4	1.7	505	4	US-09-621-976-15639	Sequence 15639, A
30	38.4	1.7	1141	4	US-09-806-708B-22	Sequence 22, Appli
31	37.8	1.7	601	4	US-09-949-016-30710	Sequence 30710, A
32	37.8	1.7	601	4	US-09-949-016-182117	Sequence 182117,
33	37.8	1.7	3190	4	US-09-949-016-5219	Sequence 5219, Ap
34	37.8	1.7	3439	4	US-09-949-016-648	Sequence 648, App
35	37.8	1.7	31602	4	US-09-949-016-16961	Sequence 16961, A
36	37.4	1.7	2718	4	US-09-614-221A-376	Sequence 376, App
37	37.2	1.7	47981	4	US-09-679-279-1	Sequence 1, Appli
38	37	1.6	832	4	US-09-621-976-2813	Sequence 2813, Ap
39	37	1.6	1059	4	US-09-248-796A-1415	Sequence 1415, Ap
40	37	1.6	1497	4	US-09-220-132-94	Sequence 94, Appl
41	36.8	1.6	1141	4	US-09-806-708B-22	Sequence 22, Appl
42	36.8	1.6	13987	2	US-08-804-227C-13	Sequence 13, Appl
43	36.8	1.6	43280	2	US-08-804-227C-1	Sequence 1, Appli
44	36.6	1.6	486	4	US-09-639-207-13	Sequence 13, Appl
45	36.4	1.6	228851	4	US-09-949-016-13781	Sequence 13781, A

ALIGNMENTS

RESULT 1
US-08-760-615-1
; Sequence 1, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; NUMBER OF INVENTIONS: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Ebola virus
; STRAIN: Zaire
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2172
; OTHER INFORMATION: /product= "Glycoprotein"

US-08-760-615-1

Query Match	66.7%	Score 1501.4	DB 3	Length 2172
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1502	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	1	ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGAATTCGAAGGACATCATTTCTTT	60	
Db	142	ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGAATTCGAAGGACATCATTTCTTT	201	
Qy	61	CTTTGGGTAAATTATCTTTTCCAAAGAAATTTTTCATTCCTCAATTCGAGTCAATTCACAAT	120	
Db	202	CTTTGGGTAAATTATCTTTTCCAAAGAAATTTTTCATTCCTCAATTCGAGTCAATTCACAAT	261	
Qy	121	AGCACATTACAGGTTAGTGATGTCGACAAACTAGTTTGTCTGACAACTGTCATCCACA	180	
Db	262	AGCACATTACAGGTTAGTGATGTCGACAAACTAGTTTGTCTGACAACTGTCATCCACA	321	
Qy	181	AATCAATTGAGATCAGTTGGAATCTCGAAGGGAATGAGTGGCAACTGAGTGCCA	240	
Db	322	AATCAATTGAGATCAGTTGGAATCTCGAAGGGAATGAGTGGCAACTGAGTGCCA	381	
Qy	241	TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCACCAAGGTGGTCAATTATGAA	300	
Db	382	TCTGCAACTAAAGATGGGGCTTCAGTCCGGTGTCCACCAAGGTGGTCAATTATGAA	441	
Qy	301	GCTGTCGAAATGGGCTGAAAACTGCTACAATCTTGAAATCAAAAACTGACGGGAGTGAG	360	
Db	442	GCTGTCGAAATGGGCTGAAAACTGCTACAATCTTGAAATCAAAAACTGACGGGAGTGAG	501	
Qy	361	TGTCCTACAGACGGCCAGACGGGAATTCGGGGCTTCCCGGTGCCGGTATGTGCAAAA	420	
Db	502	TGTCCTACAGACGGCCAGACGGGAATTCGGGGCTTCCCGGTGCCGGTATGTGCAAAA	561	
Qy	421	GTATCAGGAACGGGACCGTGTGCCGGAGACTTTTCCTTCCATAAAGAGGGTGCTTCTTC	480	
Db	562	GTATCAGGAACGGGACCGTGTGCCGGAGACTTTTCCTTCCATAAAGAGGGTGCTTCTTC	621	
Qy	481	CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC	540	
Db	622	CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC	681	
Qy	541	GTTGCAATTTCTGATCTGCCCAAGCTTAAGAGGACTTTTCAGCTCACACCCCTTGAGA	600	
Db	682	GTTGCAATTTCTGATCTGCCCAAGCTTAAGAGGACTTTTCAGCTCACACCCCTTGAGA	741	
Qy	601	GAGCCGCTCAATGCAACGGAGGACCCGTCTAGTGGCTACTATTCTACCACAATTAGATAT	660	
Db	742	GAGCCGCTCAATGCAACGGAGGACCCGTCTAGTGGCTACTATTCTACCACAATTAGATAT	801	
Qy	661	CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTCGAC	720	
Db	802	CAGGCTACCGGTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTCGAC	861	
Qy	721	TACGTCCAACTTGAATCAAGATTCCACACACAGTTTCTGCTCCAGCTGGAATGAGACAA	780	
Db	862	TACGTCCAACTTGAATCAAGATTCCACACACAGTTTCTGCTCCAGCTGGAATGAGACAA	921	
Qy	781	TATACAAAGTGGGAAAGGAGCAATACCAACGGGAACTTAATTTGGAAAGGTCAACCCCGAA	840	
Db	922	TATACAAAGTGGGAAAGGAGCAATACCAACGGGAACTTAATTTGGAAAGGTCAACCCCGAA	981	
Qy	841	ATTGATACAAATTCGGGGAGTGGGCTTCTGGGAAACTAAAAAACTCTACCTAGAAAA	900	
Db	982	ATTGATACAAATTCGGGGAGTGGGCTTCTGGGAAACTAAAAAACTCTACCTAGAAAA	1041	
Qy	901	ATTCGCAAGTGAAGAGTTGTCTTTTCCACAGTTTGTATCAAAACGGAGGCCAAAAATCAGTGGT	960	
Db	1042	ATTCGCAAGTGAAGAGTTGTCTTTTCCACAGTTTGTATCAAAACGGAGGCCAAAAATCAGTGGT	1101	
Qy	961	CAGAGTCCGGCGCAACTTCTTCGAGCCCGGACCAACCAACCACTGAAGACCAAAA	1020	
Db	1102	CAGAGTCCGGCGCAACTTCTTCGAGCCCGGACCAACCAACCACTGAAGACCAAAA	1161	

```

RESULT 2
US-09-650-086A-1
; Sequence 1, Application US/09650086A
; Patent No. 6630144
; GENERAL INFORMATION:
; APPLICANT: Hart, Mary Kate
; APPLICANT: Wilson, Julie A.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Prophylactic and Therapeutic Monoclonal Antibodies
; FILE REFERENCE: 003/198/SAP
; CURRENT APPLICATION NUMBER: US/09/650,086A
; CURRENT FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: US 60/151,505
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 1
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Ebola Zaire
US-09-650-086A-1

```

	Query Match	66.6%	Score 1499.8	DB 4	Length 2298				
	Best Local Similarity	99.9%	Pred. No. 0						
	Matches 150;	Conservative	0	Mismatches	2	Indels	0	Gaps	0
Qy	1	ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAGAGGACATCATTCCTTT	60						
Db	59	ATGGGCGTTACAGGAATATTGCAGTTACCTCGTGATCGATTCAGAGGACATCATTCCTTT	118						
Qy	61	CTTTTGGGTAAATTATCCCTTTTCCAAAGAAACATTTTCCATCCCATTTGAGTGATCCACAAT	120						
Db	119	CTTTTGGGTAAATTATCCCTTTTCCAAAGAAACATTTTCCATCCCATTTGAGTGATCCACAAT	178						
Qy	121	AGCACATTACAGGTTAGTGATGTGCAAAACTAGTTTGTGTCGACAAACTGTGATCCACA	180						

Db 179 AGCATTACAGTTAGTGTGACAACTAGTTTGTGTGACAACTGTCTATCCACA 238
Qy 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db 239 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 298
Qy 241 TCTGCAATTAAGATGGGCTTCAGGTCGGGTCCGCTCCACCAAGGTTGTCNAATATGAA 300
Db 299 TCTGCAATTAAGATGGGCTTCAGGTCGGGTCCGCTCCACCAAGGTTGTCNAATATGAA 358
Qy 301 GCTGTGTAATGGGCTGAAAACCTGTCTACAACTTTGAAATCAAAAACCTGACGGAGTGAG 360
Db 359 GCTGTGTAATGGGCTGAAAACCTGTCTACAACTTTGAAATCAAAAACCTGACGGAGTGAG 418
Qy 361 TGTCTACAGACGCCAGACGGGATTCGGGGCTTCCCCGGTGCCTGTATGTGCACAAA 420
Db 419 TGTCTACAGACGCCAGACGGGATTCGGGGCTTCCCCGGTGCCTGTATGTGCACAAA 478
Qy 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGGCTTCCATTAAGAGGGTGTCTTCTTC 480
Db 479 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGGCTTCCATTAAGAGGGTGTCTTCTTC 538
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
Db 539 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 598
Qy 541 GTTGCATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAAGCTCACACCCCTTGAGA 600
Db 599 GTTGCATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAAGCTCACACCCCTTGAGA 658
Qy 601 GAGCCGGTCAATGCAACGGAGACCGCTAGTGTGCTACTATTCTACCAATATAGATAT 660
Db 659 GAGCCGGTCAATGCAACGGAGACCGCTAGTGTGCTACTATTCTACCAATATAGATAT 718
Qy 661 CAGGCTACCGGTTTTGGAAACCAATCAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 720
Db 719 CAGGCTACCGGTTTTGGAAACCAATCAGACAGAGTACTTGTTCGAGGTTGACAATTTGACC 778
Qy 721 TACGTCCTCACTTGAATCAAGATTTACACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 779 TACGTCCTCACTTGAATCAAGATTTACACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 838
Qy 781 TATACAGTGGGAAAAGGAGCAATACCGGGAATACTAATTTGGAAGGTCAACCCCGAA 840
Db 839 TATACAGTGGGAAAAGGAGCAATACCGGGAATACTAATTTGGAAGGTCAACCCCGAA 898
Qy 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAATAAACTCACTAGAAAA 900
Db 899 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAATAAACTCACTAGAAAA 958
Qy 901 ATTCGAGTGAAGATTTGTCTTTTACAGTTGTATCAAAACGGAGCCAAAAACATCAGTGGT 960
Db 959 ATTCGAGTGAAGATTTGTCTTTTACAGTTGTATCAAAACGGAGCCAAAAACATCAGTGGT 1018
Qy 961 CAGAGTCCGGCCGGAATTTCTTCGACCCAGGGAACAAACAACTGAAGACCAACAAA 1020
Db 1019 CAGAGTCCGGCCGGAATTTCTTCGACCCAGGGAACAAACAACTGAAGACCAACAAA 1078
Qy 1021 ATCATGGCTTCAGAAATTTCTTCGAATGGTTCAAGTGCAAGTCAAGGAGGGAAGCT 1080
Db 1079 ATCATGGCTTCAGAAATTTCTTCGAATGGTTCAAGTGCAAGTCAAGGAGGGAAGCT 1138
Qy 1081 GCAGTGTGCACTTAAACACCTTGCACAAATCTCCAGAGTCCCAATCTCCACAACTC 1140
Db 1139 GCAGTGTGCACTTAAACACCTTGCACAAATCTCCAGAGTCCCAATCTCCACAACTC 1198
Qy 1141 AAACAGGTCGGGACAAAGCAACCCATATAACCCGCTGTATATACTTGACATCTCTGAG 1200
Db 1199 AAACAGGTCGGGACAAAGCAACCCATATAACCCGCTGTATATACTTGACATCTCTGAG 1258
Qy 1201 GCAACTCAAGTTGAACAACTACCGCAGAAACAGACAGGACGACGACGCTCCGACACT 1260

Db 1259 GCAACTCAAGTTGAACAACTACCGCAGAACACACAAACGACGACGCTCCGACACT 1318
Qy 1261 CCCTCTGCCACGACCGGAGCGGACCCCAAAAGAGAGAAACCAACACGAGCAAGAGC 1320
Db 1319 CCCTCTGCCACGACCGGAGCGGACCCCAAAAGAGAGAAACCAACACGAGCAAGAGC 1378
Qy 1321 ACTGACTTCTGGAGCCCGCCACCAACAACTGCCCAAAACCAACGAGGACGCTGCGC 1380
Db 1379 ACTGACTTCTGGAGCCCGCCACCAACAACTGCCCAAAACCAACGAGGACGCTGCGC 1438
Qy 1381 AACAAACACTCATCATACCAAGATACCGGAGAGAGAGTGCAGCAGCGGGAAGCTAGGC 1440
Db 1439 AACAAACACTCATCATACCAAGATACCGGAGAGAGAGTGCAGCAGCGGGAAGCTAGGC 1498
Qy 1441 TTAATTAACCAATATATTTCTGGAGTGCAGGACTGATCAAGGCGGAGAGAACTCGA 1500
Db 1499 TTAATTAACCAATATATTTCTGGAGTGCAGGACTGATCAAGGCGGAGAGAACTCGA 1558
Qy 1501 CGA 1503
Db 1559 AGA 1561

RESULT 3
US-09-336-910A-1
; Sequence 1, Application US/09336910A
; Patent No. 6517842
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan P.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143
; CURRENT APPLICATION NUMBER: US/09/336,910A
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: US 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
; ORGANISM: Marburg Virus
US-09-336-910A-1

Query Match 33.0%; Score 742.2; DB 4; Length 11460;
Best Local Similarity 98.9%; Pred. No. 7.3e-226;
Matches 747; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1492 AGAATCGACGATCGATCTCTGGAGGGAAGCGACATGTTCCCTTTCTGGATCGGTTA 1551
Db 7233 AGAAGAAAGCAAGATCTCTGGAGGGAAGCGACATGTTCCCTTTCTGGATCGGTTA 7292
Qy 1552 ATAAATGCTCCAATTTGATTTTGACCCAGTTCCAAATACAAAAACAATCTTTGATGAATCC 1611
Db 7293 ATAAATGCTCCAATTTGATTTTGACCCAGTTCCAAATACAAAAACAATCTTTGATGAATCC 7352
Qy 1612 TCTAGTTCCTGTGCTCGGCTGAGGAAGATCAACATGCTCCCAATATTTAGTTTAACT 1671
Db 7353 TCTAGTTCCTGTGCTCGGCTGAGGAAGATCAACATGCTCCCAATATTTAGTTTAACT 7412
Qy 1672 TTATCTTATTTTCTTAATATAATGAGAACTGCTCTCTGGAGAAATGAGAAATGAT 1731
Db 7413 TTATCTTATTTTCTTAATATAATGAGAACTGCTCTCTGGAGAAATGAGAAATGAT 7472
Qy 1732 TGTGATGACAGAGTTGAAGAAATTTGAGGCGTTTCAAGAGGATGACCTGGCGGAGGCTCAGT 1791
Db 7473 TGTGATGACAGAGTTGAAGAAATTTGAGGCGTTTCAAGAGGATGACCTGGCGGAGGCTCAGT 7532

```
QY 1792 TGGATACCCGTTTTTGGCCCTCGAATTGAAGACCTTTACACCTGCTGTTTAAATTTAAAAAT 1851
Db 7533 TGGATACCCGTTTTTGGCCCTCGAATTGAAGACCTTTACACCTGCTGTTTAAATTTAAAAAT 7592
QY 1852 CAAACAAATTTGGTCTGCAGGTTGAGGCTCTAGCCCAATCAAACTGCCAAATCCCTTGGAA 1911
Db 7593 CAAACAAATTTGGTCTGCAGGTTGAGGCTCTAGCCCAATCAAACTGCCAAATCCCTTGGAA 7652
QY 1912 CTCCTATTGAGAGTCACAACTGAGGAAAGAACATCTCTCTTAATCAATAGACATGCTATT 1971
Db 7653 CTCCTATTGAGAGTCACAACTGAGGAAAGAACATCTCTCTTAATCAATAGACATGCTATT 7712
QY 1972 GACTTTCTACTCACAAGATGGGAGGAAACATCAAAAGTCTTTGGACCTGATTGTTGCATC 2031
Db 7713 GACTTTCTACTCACAAGATGGGAGGAAACATCAAAAGTCTTTGGACCTGATTGTTGCATC 7772
QY 2032 GGGATAGAAGACTTGTCCAAAATATTTTCAGAGCAAAATGACCAAAATTTAAAAGACGAA 2091
Db 7773 GGGATAGAAGACTTGTCCAAAATATTTTCAGAGCAAAATGACCAAAATTTAAAAGACGAA 7832
QY 2092 CAAAAAGAGGGGACTGTTGGGCTGCGGTGGTAAATGGTGACATCCGACTGGGGTGT 2151
Db 7833 CAAAAAGAGGGGACTGTTGGGCTGCGGTGGTAAATGGTGACATCCGACTGGGGTGT 7892
QY 2152 CTTACTAACTTGGGCAATTTTGCTACTATTATTCATAGCTGTCTTGATTGCTCTATCCTGT 2211
Db 7893 CTTACTAACTTGGGCAATTTTGCTACTATTATTCATAGCTGTCTTGATTGCTCTATCCTGT 7952
QY 2212 ATTTGCTGATCTTTACTAAATATATTCGGATAACG 2246
Db 7953 ATTTGCTGATCTTTACTAAATATATTCGGATAACG 7987
```

RESULT 4

```
US-08-760-615-3
; Sequence 3, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2164 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Mueoke
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 119..2164
US-08-760-615-3

Query Match 32.9%; Score 740.2; DB 3; Length 2164;
Best Local Similarity 98.9%; Pred. No. 9.2e-226;
Matches 745; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1492 AGAATCCACCATCGATCCTCTGGAGGGAAGCGACATGTTCCCTTTCTTGATGGGTGA 1551
Db 1412 AGAAGAAAGGAAGATCCTCTGGAGGGAAGCGACATGTTCCCTTTCTTGATGGGTGA 1471
QY 1552 ATAAATGCTCCAAATTTGATTTGACCCAGTTCCAAATACAAAAACAATCTTTGATGAATCC 1611
Db 1472 ATAAATGCTCCAAATTTGATTTGACCCAGTTCCAAATACAAAAACAATCTTTGATGAATCC 1531
QY 1612 TCTAGTTCTGGTGCCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATTAGTTTAACT 1671
Db 1532 TCTAGTTCTGGTGCCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATTAGTTTAACT 1591
QY 1672 TTATCTTATTTTCCCTAAATATATAATGAGAACACTGCTCTCTGGAGAAATGAGAATGAT 1731
Db 1592 TTATCTTATTTTCCCTAAATATATAATGAGAACACTGCTCTCTGGAGAAATGAGAATGAT 1651
QY 1732 TGTGATGCAGAGTTAAAGAAATTTGAGAGCGTTTCAAGAGGATGACCTGGCCGCGAGGCTCAGT 1791
Db 1652 TGTGATGCAGAGTTAAAGAAATTTGAGAGCGTTTCAAGAGGATGACCTGGCCGCGAGGCTCAGT 1711
QY 1792 TGGATACCCGTTTTTGGCCCTGGAATTTGAAGACCTTTACATGCTGTTTAAATTTAAAAAT 1851
Db 1712 TGGATACCCGTTTTTGGCCCTGGAATTTGAAGACCTTTACATGCTGTTTAAATTTAAAAAT 1771
QY 1852 CAAACAAATTTGGTCTGCAGGTTGAGGCGTCTAGCCCAATCAAACTGCCAAATCCCTTGGAA 1911
Db 1772 CAAACAAATTTGGTCTGCAGGTTGAGGCGTCTAGCCCAATCAAACTGCCAAATCCCTTGGAA 1831
QY 1912 CTCCTATTGAGAGTCACAACTGAGGAAAGAACATTTCTCTTAATCAATAGACATGCTATT 1971
Db 1832 CTCCTATTGAGAGTCACAACTGAGGAAAGAACATTTCTCTTAATCAATAGACATGCTATT 1891
QY 1972 GACTTTCTACTCACAAGATGGGAGGAAACATGCAAAAGTCTTTGGACCTGATTGTTGCATC 2031
Db 1892 GACTTTCTACTCACAAGATGGGAGGAAACATGCAAAAGTCTTTGGACCTGATTGTTGCATC 1951
QY 2032 GGGATAGAAGACTTGTCCAAAATATTTTCAGAGCAAAATTTGACCAAAATTTAAAAGACGAA 2091
Db 1952 GGGATAGAAGACTTGTCCAAAATATTTTCAGAGCAAAATTTGACCAAAATTTAAAAGACGAA 2011
QY 2092 CAAAAAGAGGGGACTGTTGGGCTGCGGTGTAATGTTGACATCCGACTGGGGTGT 2151
Db 2012 CAAAAAGAGGGGACTGTTGGGCTGCGGTGTAATGTTGACATCCGACTGGGGTGT 2071
QY 2152 CTTACTAACTTGGGCAATTTTGCTACTATTATTCATAGCTGTCTTGATTGCTCTATCCTGT 2211
Db 2072 CTTACTAACTTGGGCAATTTTGCTACTATTATTCATAGCTGTCTTGATTGCTCTATCCTGT 2131
QY 2212 ATTTGCTGATCTTTACTAAATATATATTCGGATAA 2244
Db 2132 ATTTGCTGATCTTTACTAAATATATATTCGGATAA 2164

RESULT 5
US-08-760-615-5
; Sequence 5, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
```

APPLICANT: Haynes, Joel R
APPLICANT: Schmaljohn, Connie S
APPLICANT: Fuller, Deborah L
APPLICANT: Schmaljohn, Alan
APPLICANT: Jahrling, Peter B
TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,615
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 110229.91241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2247 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Marburg virus
STRAIN: Ravn
FEATURE:
NAME/KEY: CDS
LOCATION: 97..2142
US-08-760-615-5

Query Match 24.2%; Score 545.2; DB 3; Length 2247;
Best Local Similarity 82.5%; Pred. No. 2.8e-163; Mismatches 133; Indels. 0; Gaps 0;
Matches 625; Conservative 0;

```
QY 1492 AGAAGCTCGACGATCGATCTCTGGAGGGAAGCGACATGTTCCCTTTTCTGGATGGGTTA 1551
DB 1390 AGAAGAAACGAAGCATTTCTGGAAGAGGTGATATATCCCGTTTCTAGATGGGTTA 1449

QY 1552 ATAAATGCTCCAAATGATTGACCGAGTTCCAAATACAAAACAAATCTTTGATGAATCC 1611
DB 1450 ATAAATGCTGAATGATTTGATCCAAATCCAAACACAGAAACAAATCTTTGATGAATCT 1509

QY 1612 TCTAGTCTGGTCTGGCTGAGGAGATCAACATGCTCCCGCCCAATATTAGTTTAACT 1671
DB 1510 CCCAGCTTTAACTTCAACTAATGAGGAACACACATCTCCCGCCCAATATTAGTTTAACT 1569

QY 1672 TTATCTTATTTTCCATAATAATGAGAACACTGCTACTCTGGAGAAAATGAGAAATGAT 1731
DB 1570 TTCTCTTATTTTCCGATATAAATGAGAGTACTGCTACTCTGGGAAAAACGAGATGAT 1629

QY 1732 TGTGATGAGAGTTAAGAAATTTGAGCGGTTTCAGGAGGATGACCTGGCGCGAGGCTCAGT 1791
DB 1630 TGTGATGAGAGTTGAGGATTTGAGGTGTGAGGAGGACGATTTGGCGCGAGGCTTAGC 1689

QY 1792 TCGATACCGTTTTTGGCCCTGGATTTAGGACATTTACAGCTGTTTAAATTAATAAAT 1851
DB 1690 TGGATACCAATTTTTTGGCCCTGGATTCGAATCGAAGGACTCTATACTGCGCGGTTTAAATCAAAAT 1749
```

```
QY 1852 CAAACCAATTTGGTCTCAGGTTGAGCGCTAGCCCAATCAAACTGCCAAATCCTTGGAA 1911
DB 1750 CAGAACCAATTTAGTTTCTAGGTTGAGCGCTAGCTAATCAAACTGCTAAATCCTTGGAG 1809

QY 1912 CTCTTATTGAGGTCACAACTGAGGAAGAACATTTCTTCTTAATCAATAGACATGCTATT 1971
DB 1810 CTCTTGTGTAAGGGTCACAACTGAGGAAGAACATTTTCTTCTTAATCAATAGGATGCAATT 1869

QY 1972 GACTTTCTACTCACAAGATGGGAGGAACATGCAAACTGCTTGGACCTGATTGTTGCAATC 2031
DB 1870 GACTTTTGTCTAGAGGTGGGCGGAACATGCAAGGTGCTAGGACCTGATTGTTGCAATC 1929

QY 2032 GGGATAGAAGACTTGTCCAAAATATTTCAGAGCAAAATTCAGCAAAATTAAGAGAGCAAA 2091
DB 1930 GGAATAGAAGATCTATCTAAAAATATCTCAGAACAAATCGACAAAATCAGAAAGGATGAA 1989

QY 2092 CAAAAAGAGGGGACTGTTGGGCTCTGGGTGCTGTAATGTTGGACATCGGACTGGGCTGTT 2151
DB 1990 CAAAAAGAGGAAACTGGCTGGGGTCTAGGTGGCAAAATGGTGGACATCTGACTGGGGTGT 2049

QY 2152 CTACTAACTTGGGCAATTTTGTCTACTATTATCCATAGCTGTCTTGAATGCTCTATCTCTGT 2211
DB 2050 CTCACCAATTTGGGCACTCTGCTACTATTATCTATAGCTGTCTTGAATGCTCTCTCTGT 2109

QY 2212 ATTTGCTGTATCTTTTACTAAATATATCGGATAACGAA 2249
DB 2110 ATCTGCTGTATCTTCACTAAATATATCGGATGACATAA 2147
```

RESULT 6

US-09-621-976-8976
; Sequence 8976, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8976
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8976

```
Query Match 2.2%; Score 49.2; DB 4; Length 399;  
Best Local Similarity 12.2%; Pred. No. 6.5e-05;  
Matches 41; Conservative 163; Mismatches 131; Indels 1; Gaps 1;  
QY 899 AAATTCGAGTGAAGAGTTGTTTTCACAGTTGTATCAACAGGAGCCAAAACATCAGTG 958  
DB 20 RAYGMYTRGSKSWRAKMSWMKRRRRRRRAWMWMSKMWKSKSWRSWRSWMTKRMKG 79  
QY 959 GTCAGAGTCCGGCCGGAATCTTCCGACCAGGAGCAACACACAACTGAGAACCA 1018  
DB 80 RGAASWAGYMSWMTYMTTTRWRYRYRKACCTKWRAGAGWAGWAWAYAKWYMAWHRRT 139  
QY 1019 AAATCATGCGCTTCAGAAAATTCCTCTGCAATGGTTTCAAG-TGCACAGTCAAGGAGGAA 1077  
DB 140 AMKWAMWMSKRSRRRRRAWYAWMYMARTMTMGMBRASCYRGAYMASAGYMYWYMM 199  
QY 1078 GCTCAGTGTGCGATCTAAACAAACCTTGGCCCAATCTCCAGAGTCCCAATCCCTCACA 1137  
DB 200 RKNWYAGWMSMRKWTTRRCASYSWSSYCNWGAAMWYMTKTSRWSYSSYRCTKYRRS 259  
QY 1138 ACCAAACAGGTCGGGACACAGACCCCAATATACACCGGTGTATTAACCTTGACATCTCT 1197  
DB 260 CCCWMSMCYKWTYRYSYSCSYSYKTKRASCCCMCCMKRKMWMMWMMWMMWMMWTKCT 319
```



```
QY 1336 CCGGCCACCAACAGTCTCCCAAAACACAGCGAGACCGCTGGCAACAAACAC 1391
Db 98565 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 98620

RESULT 10
US-09-949-016-13845/c
; Sequence 13845, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13845
; LENGTH: 152132
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13845

Query Match 2.0%; Score 44.8; DB 4; Length 152132;
Best Local Similarity 47.0%; Pred. No. 0.14;
Matches 139; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 1096 ACAACCTTGGCAGATCTCCAGAGTCCCAATCCCTCACAACCAAAACAGGTCGGGAC 1155
Db 136812 ACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136753

QY 1156 AACAGCACCATAATACACCCGCTGTATAAATTTGACATCTCTGAGGCAACTCAAGTTGAA 1215
Db 136752 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136693

QY 1216 CAACATCACCAGCAAGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1275
Db 136692 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136633

QY 1276 GCAGCCGACCCCAAAAGCAGAGCAACCAACCAACCAACCAACCAACCAACCAAC 1335
Db 136632 ACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136573

QY 1336 CCGGCCACCAACAGTCTCCCAAAACACAGCGAGACCGCTGGCAACAAACAC 1391
Db 136572 ACCACCACTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136517

RESULT 11
US-09-949-016-12371/c
; Sequence 12371, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12371
; LENGTH: 152145
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12371

Query Match 2.0%; Score 44.8; DB 4; Length 152145;
Best Local Similarity 47.0%; Pred. No. 0.14;
Matches 139; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 1096 ACAACCTTGGCAGATCTCCAGAGTCCCAATCCCTCACAACCAAAACAGGTCGGGAC 1155
Db 136812 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136753

QY 1156 AACAGCACCATAATACACCCGCTGTATAAATTTGACATCTCTGAGGCAACTCAAGTTGAA 1215
Db 136752 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136693

QY 1216 CAACATCACCAGCAAGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1275
Db 136692 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136633

QY 1276 GCAGCCGACCCCAAAAGCAGAGCAACCAACCAACCAACCAACCAACCAACCAAC 1335
Db 136632 ACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136573

QY 1336 CCGGCCACCAACAGTCTCCCAAAACACAGCGAGACCGCTGGCAACAAACAC 1391
Db 136572 ACCACCACTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 136517

RESULT 12
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOMLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
```

```
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7218 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ CLONE: pTZ9pt-Fls
/ US-08-232-463-14

Query Match          2.0%; Score 44.4; DB 1; Length 7218;
Best Local Similarity 8.7%; Pred. No. 0.019;
Matches 39; Conservative 209; Mismatches 200; Indels 0; Gaps 0;

QY 629 CTAGTGGCTACTATTCTACCAATAGATATCAGGCTACCGGTTTGGAAACCAATGAGA 688
Db 1466 CAAGTAGTTAAAGAGATAGAAGAAATTTGGTACRNRNRNRNRNRNRNRNRNRNR 1407
QY 689 CAGAGTACTGTTTCGAGGTTGCAATTTGACCTACGTCCTCAATTTCAATCAAGATTCACAC 748
Db 1406 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1347
QY 749 CACAGTTTCTGCTCCAGCTGAATGAGACAAATATATACAGTGGGAAAGGACGAATACCA 808
Db 1346 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1287
QY 809 CGGGAAACCTAATTTGGAAAGTCAACCCCGAAATTTGATACAAATCGGGGAGTGGGCT 868
Db 1286 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1227
QY 869 TCTGGGAACTAAATAAAACCTCCTAGAGAAATTCGAGTGAAGAGTTGCTTTTCACAG 928
Db 1226 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1167
QY 929 TTGTATCAACGAGCGCAAAACATCAGTGGTCAGAGTCCGGCGGCAACTTCTCCGACC 988
Db 1166 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1107
QY 989 CAGGACCAACCAACCACTGAAGACCAACAAATCATGCTTCAGAAAATTCCTCTGCAA 1048
Db 1106 RRRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNRNR 1047
QY 1049 TGGTTCAAGTCAGCTCAAGGAGGGA 1076
Db 1046 GCAGCAAGCTCGGAATTAATCTGTGA 1019

RESULT 13
US-09-949-016-14613/c
; Sequence 14613, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14613
; LENGTH: 30656
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(30656)
; OTHER INFORMATION: n = A,T,C or G
```

```
US-09-949-016-14613
Query Match          2.0%; Score 44.2; DB 4; Length 30656;
Best Local Similarity 47.1%; Pred. No. 0.065;
Matches 136; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 1099 ACCCTTGGCAACAATCTCCAGAGTCCCAATCCCTCAACAACCAACGAGTCCGGACAAAC 1158
Db 3814 ACCATCACCACCAACCAATCACCAGCAACCAACCAACCAACCAACCAACCAACCA 3755
QY 1159 AGCACCCATATACACCGTGTATATAAAGTTTGACATCTCTGAGGCAACTCAAGTTGAACA 1218
Db 3754 ATCAGAGCACCAATCACCAGCAACCAATACCACTACCACTACCACTACCACTACCA 3695
QY 1219 CATCACCGCAGAACAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1278
Db 3694 ACCAGCACCAACCAATCACAACCAACCACTACCACTACCACTACCACTACCACTAC 3635
QY 1279 GCGGACCCCCCAAGAGAGAAACCAACGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1338
Db 3634 ACCATTACCACTACTTACCACCAATCACCACCAATCACCAGCACCAGCACCACCAACC 3575
QY 1339 GCGACCAACAAGTCCCAAAACCAACGAGCGGCTGGCAACAACA 1387
Db 3574 ATCAACCACTTACCACCAATCACCACCACTACCACTACCACTACCACTACCACTAC 3526

RESULT 14
US-09-949-016-103758/c
; Sequence 103758, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103758
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-103758

Query Match          1.9%; Score 43.8; DB 4; Length 522;
Best Local Similarity 46.7%; Pred. No. 0.0043;
Matches 135; Conservative 1; Mismatches 153; Indels 0; Gaps 0;

QY 1099 ACCCTTGGCAACAATCTCCAGAGTCCCAATCCCTCAACAACCAACGAGTCCGGACAAAC 1158
Db 489 ACCATCACCACCAACCAACCAATCACCAGCAACCAACCAACCAACCAACCAACCA 430
QY 1159 AGCACCCATATACACCGTGTATATAAAGTTTGACATCTCTGAGGCAACTCAAGTTGAACA 1218
Db 429 ATCACCAGCACCAATCACCAGCAACCAATCACCACCACTACCACTACCACTACCACTAC 370
QY 1219 CATCACCGCAGAACAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1278
Db 369 ACCAGCACCAACCAATCACAACCAACCACTACCACTACCACTACCACTACCACTACCA 310
QY 1279 GCGGACCCCCCAAGAGAGAAACCAACGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1338
Db 309 ACCATTACCACTACTTACCACCAATCACCACCAATCACCAGCACCAGCACCACCAACC 250
QY 1339 GCGACCAACAAGTCCCAAAACCAACGAGCGGCTGGCAACAACA 1387
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 16:57:35 ; Search time 1415.31 Seconds
(without alignments)
10298.647 Million cell updates/sec

Title: US-10-066-506A-1
Perfect score: 2252
Sequence: 1 atgggcgttacaggaattt.....tatatcgataacggaattc 2252

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2252	100.0	2252	15	US-10-066-506A-1
2	1506	66.9	2039	15	US-10-066-506A-11
3	1501.4	66.7	2298	18	US-10-226-795-1
4	1501.4	66.7	6987	20	US-10-491-121-6
5	1501.4	66.7	6914	20	US-10-491-121-9
6	1501.4	66.7	7044	20	US-10-491-121-7
7	1501.4	66.7	7106	20	US-10-491-121-8

8	1501.4	66.7	7154	20	US-10-491-121-1	Sequence 1, Appli
9	1501.4	66.7	7188	20	US-10-491-121-2	Sequence 2, Appli
10	1501.4	66.7	7285	22	US-10-860-878-3	Sequence 3, Appli
11	1501.4	66.7	8199	20	US-10-491-121-29	Sequence 25, Appli
12	1501.4	66.7	8439	20	US-10-491-121-28	Sequence 28, Appli
13	1501.4	66.7	10783	20	US-10-491-121-25	Sequence 25, Appli
14	1499.8	66.6	2298	9	US-09-337-946A-1	Sequence 1, Appli
15	1499.8	66.6	2298	17	US-10-384-976-1	Sequence 1, Appli
16	1499.8	66.6	2298	22	US-10-866-833-1	Sequence 1, Appli
17	1489.4	66.1	7272	22	US-10-860-878-4	Sequence 4, Appli
18	1489.4	66.1	18959	17	US-10-353-856-19	Sequence 19, Appli
19	1489.4	66.1	18959	17	US-10-353-856-37	Sequence 37, Appli
20	1489.4	66.1	18959	17	US-10-353-856-46	Sequence 46, Appli
21	1480	65.7	6724	20	US-10-491-121-5	Sequence 5, Appli
22	1446.2	64.2	2030	21	US-10-811-353-4	Sequence 4, Appli
23	906.4	40.2	6624	20	US-10-491-121-3	Sequence 3, Appli
24	905.4	40.2	6561	20	US-10-491-121-4	Sequence 4, Appli
25	900	40.0	6324	20	US-10-491-121-37	Sequence 37, Appli
26	753	33.4	2051	15	US-10-066-506A-9	Sequence 9, Appli
27	746.6	33.2	2046	15	US-10-066-506A-7	Sequence 7, Appli
28	742.2	33.0	7778	20	US-10-491-121-30	Sequence 30, Appli
29	742.2	33.0	11460	16	US-10-267-322-1	Sequence 1, Appli
30	685.8	30.5	2046	21	US-10-811-353-2	Sequence 2, Appli
31	685.8	30.5	19112	17	US-10-353-856-27	Sequence 27, Appli
32	675.2	30.0	6467	20	US-10-491-121-10	Sequence 10, Appli
33	644.2	28.6	7005	20	US-10-491-121-31	Sequence 31, Appli
34	644.2	28.6	8256	20	US-10-491-121-32	Sequence 32, Appli
35	548.2	24.3	2046	15	US-10-066-506A-13	Sequence 13, Appli
36	546.6	24.3	2046	15	US-10-066-506A-5	Sequence 5, Appli
37	546.2	24.3	6885	20	US-10-491-121-18	Sequence 18, Appli
38	546.2	24.3	7002	20	US-10-491-121-16	Sequence 16, Appli
39	546.2	24.3	7023	20	US-10-491-121-21	Sequence 21, Appli
40	546.2	24.3	7036	20	US-10-491-121-17	Sequence 17, Appli
41	534.2	23.7	6889	20	US-10-491-121-19	Sequence 19, Appli
42	534.2	23.7	7003	22	US-10-860-878-1	Sequence 1, Appli
43	534.2	23.7	8146	20	US-10-491-121-20	Sequence 20, Appli
44	466.2	20.7	18890	17	US-10-353-856-9	Sequence 9, Appli
45	464.6	20.6	2360	17	US-10-397-635-12	Sequence 12, Appli

ALIGNMENTS

RESULT 1
US-10-066-506A-1
; Sequence 1, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 1
; LENGTH: 2252
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Ebola virus Zaire Mayinga strain Glycoprotein 2 and Marburg virus strain Musoke Glycoprotein 2
US-10-066-506A-1

Query Match 100.0%; Score 2252; DB 15; Length 2252;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGCGGTACAGGAATTGCTGATCGATTCAGAGGACATCTCTTT 60

|||||
1 ATGGCGTTACAGGAATATTGCGAGTTACCTCGTGATCGAATTCAGAGGACATCATTTCTTT 60
QY
61 CTTTGGGTAAATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
Db
61 CTTTGGGTAAATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120
QY
121 AGCACATTACAGGTTAGTGATGTCGACAAATAGTTTGTGCGACAAACTGTCATCCACA 180
Db
121 AGCACATTACAGGTTAGTGATGTCGACAAATAGTTTGTGCGACAAACTGTCATCCACA 180
QY
181 AATCAATTGAGATCAGTTGGAATCTGAAAGGGAATCGAAGGGAATGGAAGTGGTCAATTATGAA 240
Db
181 AATCAATTGAGATCAGTTGGAATCTGAAAGGGAATCGAAGGGAATGGAAGTGGTCAATTATGAA 240
QY
241 TCTGCAACTTAAAGATGGGGCTTCAGGTCGGGTGTCACCAACCAAGGTGGTCAATTATGAA 300
Db
241 TCTGCAACTTAAAGATGGGGCTTCAGGTCGGGTGTCACCAACCAAGGTGGTCAATTATGAA 300
QY
301 GCTGTGAAATGGGCTGAAAATCTGCTACAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
Db
301 GCTGTGAAATGGGCTGAAAATCTGCTACAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
QY
361 TGTCTACCAAGCGGCGCAGACGGGATTCGGGGTTCCCGGTGCGGTGATGTGCACAAA 420
Db
361 TGTCTACCAAGCGGCGCAGACGGGATTCGGGGTTCCCGGTGCGGTGATGTGCACAAA 420
QY
421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGGCTTCCATTAAGAGGGTCTTTCTTC 480
Db
421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGGCTTCCATTAAGAGGGTCTTTCTTC 480
QY
481 CTGTATGATCGACTTCTGTTCCACAGTTATCTACCGAGGAACGACTTTTCCTGAAAGTGTC 540
Db
481 CTGTATGATCGACTTCTGTTCCACAGTTATCTACCGAGGAACGACTTTTCCTGAAAGTGTC 540
QY
541 GTTGCATTTCTGATATCGCCCCAAGCTAAGAAAGGACTTCTCAGCTCACACCCCTTGAGA 600
Db
541 GTTGCATTTCTGATATCGCCCCAAGCTAAGAAAGGACTTCTCAGCTCACACCCCTTGAGA 600
QY
601 GAGCCGGTCAATGCAACGGAGGACCGCTCTAGTGGCTACTATTTACCACAATTAGATAT 660
Db
601 GAGCCGGTCAATGCAACGGAGGACCGCTCTAGTGGCTACTATTTACCACAATTAGATAT 660
QY
661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTGTTGAGGTTGACAAATTTGACC 720
Db
661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTGTTGAGGTTGACAAATTTGACC 720
QY
721 TAGCTCCAACTTGAATCAAGATTACACCAAGTTCTGCTCCAGCTGAATGAGACAATA 780
Db
721 TAGCTCCAACTTGAATCAAGATTACACCAAGTTCTGCTCCAGCTGAATGAGACAATA 780
QY
781 TATACAAGTGGGAAAGGAGCAATACCGGGAAACTAATTTGGAAGGTCAACCCCGAA 840
Db
781 TATACAAGTGGGAAAGGAGCAATACCGGGAAACTAATTTGGAAGGTCAACCCCGAA 840
QY
841 ATTGATACAACAACTCGGGAGTGGGCTTCTGGGAACTTAAAAAAACCTCACTAGAAAA 900
Db
841 ATTGATACAACAACTCGGGAGTGGGCTTCTGGGAACTTAAAAAAACCTCACTAGAAAA 900
QY
901 ATTTCGAGTGAAGAGTTGTCTTTACAGTTGTATCAAAACGGAGCCAAAAACATCAAGTGT 960
Db
901 ATTTCGAGTGAAGAGTTGTCTTTACAGTTGTATCAAAACGGAGCCAAAAACATCAAGTGT 960
QY
961 CAGAGTCCGGCGGAACTTCTTCCGACCCAGGGACCAACACACACTGAAAGCACACAAA 1020
Db
961 CAGAGTCCGGCGGAACTTCTTCCGACCCAGGGACCAACACACACTGAAAGCACACAAA 1020
QY
1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAGGCT 1080
Db
1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAGGCT 1080
QY
1081 GCAGTGTGCAATCTAAACACCCCTTGCCACAATCTCCACGAGTCCCCCAATCCCTCAACCC 1140
|||||

Db
1081 GCAGTGTGCAATCTAAACACCCCTTGCCACAATCTCCACGAGTCCCCCAATCCCTCAACCC 1140
QY
1141 AAACAGAGTCCGGCAACACAGCACCCATATAACACCCGTGTATAAACTTTGACATCTCTGAG 1200
Db
1141 AAACAGAGTCCGGCAACACAGCACCCATATAACACCCGTGTATAAACTTTGACATCTCTGAG 1200
QY
1201 GCAACTCAAGTTGAACAACTATCAGCGAGAAACAGACAAACGACAGCAGCCTCCGACACT 1260
Db
1201 GCAACTCAAGTTGAACAACTATCAGCGAGAAACAGACAAACGACAGCAGCCTCCGACACT 1260
QY
1261 CCCTCTGCCACGACCGGAGCCGACCCCAAAAGCAGAGAACCAACACGAGCAGAGAGC 1320
Db
1261 CCCTCTGCCACGACCGGAGCCGACCCCAAAAGCAGAGAACCAACACGAGCAGAGAGC 1320
QY
1321 ACTGACTTCTGGAGCCCGCCACACAAAGTCTCCCAAAACCAACAGCAGCAGACCCCTGGC 1380
Db
1321 ACTGACTTCTGGAGCCCGCCACACAAAGTCTCCCAAAACCAACAGCAGCAGACCCCTGGC 1380
QY
1381 AACAAACAACTCATCTACCAAGATAACCGAGAAAGAGTGCAGCAGCGGGAAGCTAGGC 1440
Db
1381 AACAAACAACTCATCTACCAAGATAACCGAGAAAGAGTGCAGCAGCGGGAAGCTAGGC 1440
QY
1441 TTAATTACCAATACTATTGCTGGAGTCCAGAGCTGATCAGACGCGGAGAGAACTCGA 1500
Db
1441 TTAATTACCAATACTATTGCTGGAGTCCAGAGCTGATCAGACGCGGAGAGAACTCGA 1500
QY
1501 CGATCGATCTCTGGAGGAAAGGCGACATGTTCCCTTTCTGGATGGGTAAATAATGCT 1560
Db
1501 CGATCGATCTCTGGAGGAAAGGCGACATGTTCCCTTTCTGGATGGGTAAATAATGCT 1560
QY
1561 CCAATTGATTTTGACCCAGTTCCAAATAACAAACAACTTTTGATGAATCTCTAGTTCT 1620
Db
1561 CCAATTGATTTTGACCCAGTTCCAAATAACAAACAACTTTTGATGAATCTCTAGTTCT 1620
QY
1621 GGTGCTCGCTGAGGAAAGATCAATGCTCTCCCAATATAGTTTAACTTTATCTTAT 1680
Db
1621 GGTGCTCGCTGAGGAAAGATCAATGCTCTCCCAATATAGTTTAACTTTATCTTAT 1680
QY
1681 TTTCTTAATATAAATGAGAACACTGCTTACTCTGGAGAAATAGAAATGATGTGATGCA 1740
Db
1681 TTTCTTAATATAAATGAGAACACTGCTTACTCTGGAGAAATAGAAATGATGTGATGCA 1740
QY
1741 GAGTTAAGAAATTTGAGCGGTTGAGGAGGATGACCTGCGCGCAGGAGCTCAGTTGGATACCG 1800
Db
1741 GAGTTAAGAAATTTGAGCGGTTGAGGAGGATGACCTGCGCGCAGGAGCTCAGTTGGATACCG 1800
QY
1801 TTTTGTGCCCCCTGGAATTTGAAGGACTTTTACACTGCTGTTTAAATTAATAACCAACAA 1860
Db
1801 TTTTGTGCCCCCTGGAATTTGAAGGACTTTTACACTGCTGTTTAAATTAATAACCAACAA 1860
QY
1861 TTGCTCTGAGGTTGAGGCGCTTAGCCCAATCAAACTGCAAAATCTTGGAACTCTTATTG 1920
Db
1861 TTGCTCTGAGGTTGAGGCGCTTAGCCCAATCAAACTGCAAAATCTTGGAACTCTTATTG 1920
QY
1921 AGAGTCACAACTGAGGAAAGAACATTTCTTCTTAATCAATAGACATGCTATTGACTTTCTA 1980
Db
1921 AGAGTCACAACTGAGGAAAGAACATTTCTTCTTAATCAATAGACATGCTATTGACTTTCTA 1980
QY
1981 CTCACAAGATGGGAGGAAACATGCAAACTGCTTGGACCTGATTGTCATCGGATAGAA 2040
Db
1981 CTCACAAGATGGGAGGAAACATGCAAACTGCTTGGACCTGATTGTCATCGGATAGAA 2040
QY
2041 GACTTGTCCAAAAATTAATTTTCAGAGCAAAATGACCAAAATTAATAAGGAGCAACAAAGAG 2100
Db
2041 GACTTGTCCAAAAATTAATTTTCAGAGCAAAATGACCAAAATTAATAAGGAGCAACAAAGAG 2100
QY
2101 GGGACTGTTGGGGTCTGGGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAAC 2160
Db
2101 GGGACTGTTGGGGTCTGGGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAAC 2160
QY
2161 TTGGGCATTTTGTCTACTATTATTCATAGTGTCTTGAATGCTCTATCTCTGATTTTGTCTG 2220
Db
2161 TTGGGCATTTTGTCTACTATTATTCATAGTGTCTTGAATGCTCTATCTCTGATTTTGTCTG 2220

QY 2221 ATCTTTACTAAATATATATCGGATAACGGAATTC 2252
|||||
Db 2221 ATCTTTACTAAATATATCGGATAACGGAATTC 2252
|||||

RESULT 2

US-10-066-506a-11
; Sequence 11, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 11
; LENGTH: 2039
; TYPE: DNA
; ORGANISM: Ebola virus Zaire strain
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Ebola virus Glycoprotein 1 and Ebola vi
; OTHER INFORMATION: Glycoprotein 2
US-10-066-506a-11

Query Match 66.9%; Score 1506; DB 15; Length 2039;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGCGGTTACAGGAATATGCGATTACCTGTCGATCGATTCAAGAGGACATCTCTTT 60
|||||
Db 1 ATGGCGGTTACAGGAATATGCGATTACCTGTCGATCGATTCAAGAGGACATCTCTTT 60
|||||
QY 61 CTTTGGGTAATTTATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCAAAT 120
|||||
Db 61 CTTTGGGTAATTTATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCAAAT 120
|||||
QY 121 AGCATTACAGGTTAGTGTGCAAACTAGTTTGTGTCGATCAAACTGTGATCAACA 180
|||||
Db 121 AGCATTACAGGTTAGTGTGCAAACTAGTTTGTGTCGATCAAACTGTGATCAACA 180
|||||
QY 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCA 240
|||||
Db 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCA 240
|||||
QY 241 TCTGCAACTAAAGATGGGCTTTCAGGTCCGGTGTCCCAACAAAGGTGGTCAATTTATGAA 300
|||||
Db 241 TCTGCAACTAAAGATGGGCTTTCAGGTCCGGTGTCCCAACAAAGGTGGTCAATTTATGAA 300
|||||
QY 301 GCTGTGTAATGGCTGAAAACTGCTCAATCTTGAAATCAAAAACTTGAAGAGGTGAG 360
|||||
Db 301 GCTGTGTAATGGCTGAAAACTGCTCAATCTTGAAATCAAAAACTTGAAGAGGTGAG 360
|||||
QY 361 TGTCTACAGCAGCGCCAGAGCGGATTCGGGGCTTCCCGGTGCGCGGTATGTCACAAA 420
|||||
Db 361 TGTCTACAGCAGCGCCAGAGCGGATTCGGGGCTTCCCGGTGCGCGGTATGTCACAAA 420
|||||
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATTAAGAGGGTCTTTCTTC 480
|||||
Db 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATTAAGAGGGTCTTTCTTC 480
|||||
QY 481 CTGTATGATCGACTTGTCTTCCAGTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
|||||
Db 481 CTGTATGATCGACTTGTCTTCCAGTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
|||||
QY 541 GTTGCATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCAGCTCACCCCTTGAGA 600
|||||

RESULT 3
US-10-226-795-1
; Sequence 1, Application US/10226795
; Publication No. US20040053865A1
; GENERAL INFORMATION:
; APPLICANT: HART, MARY KATE

Db 541 GTTGCAATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTTCAGCTCACACCCCTTGAGA 600
QY 601 GAGCCGCTCAATGCAACGAGAGCCCGCTAGTGGCTACTATTCTTCAACCAATTAGATAT 660
Db 601 GAGCCGCTCAATGCAACGAGAGCCCGCTAGTGGCTACTATTCTTCAACCAATTAGATAT 660
QY 661 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTGTTCGAGGTTGCAATTTGACC 720
Db 661 CAGGCTACCGGTTTGGAAACCAATGACAGAGTACTTGTTCGAGGTTGCAATTTGACC 720
QY 721 TACGTCCAACTTGAATCAAGATTCAACACCAAGTTTCTGCTCCAGCTGGAATGAGACAATA 780
Db 721 TACGTCCAACTTGAATCAAGATTCAACACCAAGTTTCTGCTCCAGCTGGAATGAGACAATA 780
QY 781 TATACAAGTGGGAAAAGAGGCAATATACACGGAATAAATAATTTGGAAGGTCAACCCGAA 840
Db 781 TATACAAGTGGGAAAAGAGGCAATATACACGGAATAAATAATTTGGAAGGTCAACCCGAA 840
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAAATAAAAAACCTCAGTAGAAA 900
Db 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAAAATAAAAAACCTCAGTAGAAA 900
QY 901 ATTCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAAAACGAGGCCAAAAACATCAGTGT 960
Db 901 ATTCGAGTGAAGAGTTGTCTTTTCAAGTTGTATCAAAACGAGGCCAAAAACATCAGTGT 960
QY 961 CAGAGTCCGGCGGAACTTTTCCGACCCAGGACCAACAACTGAAGAGACACAAA 1020
Db 961 CAGAGTCCGGCGGAACTTTTCCGACCCAGGACCAACAACTGAAGAGACACAAA 1020
QY 1021 ATCATGGCTTCAGAAAATTTCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGAGCT 1080
Db 1021 ATCATGGCTTCAGAAAATTTCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGAGCT 1080
QY 1081 GCAGTGTGCAATCTTAAACAACTTGGCCCAATCTTCCAGAGTCCCAATCCTCACAACC 1140
Db 1081 GCAGTGTGCAATCTTAAACAACTTGGCCCAATCTTCCAGAGTCCCAATCCTCACAACC 1140
QY 1141 AAACAGGTCCGGACAAAGCACCATAATACACCGGTGTATAAACTTGAATCTCTGAG 1200
Db 1141 AAACAGGTCCGGACAAAGCACCATAATACACCGGTGTATAAACTTGAATCTCTGAG 1200
QY 1201 GCAACTCAAGTTGAACACATCTACCGCAGACAGACAGCAGCAGCAGCTCCGACACT 1260
Db 1201 GCAACTCAAGTTGAACACATCTACCGCAGACAGACAGCAGCAGCAGCTCCGACACT 1260
QY 1261 CCCTCTGCACGCGCAGCGGACCCCAAAAGCAGAGAACCAACACAGCAAGAGC 1320
Db 1261 CCCTCTGCACGCGCAGCGGACCCCAAAAGCAGAGAACCAACACAGCAAGAGC 1320
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAACCAACAGCGAGACCGCTGGC 1380
Db 1321 ACTGACTTCTGGACCCCGCCACCAACAAAGTCCCAACCAACAGCGAGACCGCTGGC 1380
QY 1381 AAGAAACACTCATCAACCAAGTACCGGAGAGAGAGTCCAGCAGCGGAGAGCTAGGC 1440
Db 1381 AAGAAACACTCATCAACCAAGTACCGGAGAGAGAGTCCAGCAGCGGAGAGCTAGGC 1440
QY 1441 TTAATTACCAATCTATTGTGAGTGCAGGACTGATCAAGCGGAGAGAACTCGA 1500
Db 1441 TTAATTACCAATCTATTGTGAGTGCAGGACTGATCAAGCGGAGAGAACTCGA 1500
QY 1501 CGATCG 1506
Db 1501 CGATCG 1506

```
; APPLICANT: WILSON, JULIE
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND COMPLEMENTARITY-DETERMINING
; FILE REFERENCE: ARMY 166
; CURRENT APPLICATION NUMBER: US/10/226,795
; CURRENT FILING DATE: 2002-11-18
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Ebola virus
US-10-226-795-1

Query Match      66.7%; Score 1501.4; DB 18; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGTTACAGGAATATTGCAATTGACGTTACCTCGTGATCGAATTCAGAGGACATCATTTCTTT 60
DB 59 ATGGGCGTTACAGGAATATTGCAATTGACGTTACCTCGTGATCGAATTCAGAGGACATCATTTCTTT 118
QY 61 CTTTGGGTAATATCTCTTTTCCAAAGAACATTTTCATCCCACTTGGAGTCAATCCACAAT 120
DB 119 CTTTGGGTAATATCTCTTTTCCAAAGAACATTTTCATCCCACTTGGAGTCAATCCACAAT 178
QY 121 AGCACATTACAGGTAGTGTGATGTCGACAACTAGTTTGTGTCGACAACTGTCATCCACA 180
DB 179 AGCACATTACAGGTAGTGTGATGTCGACAACTAGTTTGTGTCGACAACTGTCATCCACA 238
QY 181 AATCAATTGAGATCAGTTGGATGATCTCGAAGGGAATGAGTGGAAGTGGCACTGACGTGCCA 240
DB 239 AATCAATTGAGATCAGTTGGATGATCTCGAAGGGAATGAGTGGAAGTGGCACTGACGTGCCA 298
QY 241 TCTGCACTAAAGATGGGGCTTCAGGTCGGGTGTCACCAAGGTGCTCAATATGAA 300
DB 299 TCTGCACTAAAGATGGGGCTTCAGGTCGGGTGTCACCAAGGTGCTCAATATGAA 358
QY 301 GCTGGTGAATGGCTGAAACTGCTACAACTTTGAAATCAAAAACCTGACGGGAGTGAG 360
DB 359 GCTGGTGAATGGCTGAAACTGCTACAACTTTGAAATCAAAAACCTGACGGGAGTGAG 418
QY 361 TGTCTACAGACGCGCAGACGGGATTCGGGGCTTCCCGCGTTCGGGTATGTGCACAAA 420
DB 419 TGTCTACAGACGCGCAGACGGGATTCGGGGCTTCCCGCGTTCGGGTATGTGCACAAA 478
QY 421 GTATCAGGAAACGGGACCGTGTGCGAGACTTTGCTTCATTAAGAGGGTCTTCTTC 480
DB 479 GTATCAGGAAACGGGACCGTGTGCGAGACTTTGCTTCATTAAGAGGGTCTTCTTC 538
QY 481 CTGTATGATCGACTTCTTCCACAGTTATCTACCGAGGAACGACTTTGCTCAAGGTGTC 540
DB 539 CTGTATGATCGACTTCTTCCACAGTTATCTACCGAGGAACGACTTTGCTCAAGGTGTC 598
QY 541 GTTGCAATTTCTGATATCGCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 600
DB 599 GTTGCAATTTCTGATATCGCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 658
QY 601 GAGCCGGTCAATGCAACGGAGACCCGCTAGTGGCTACTATTCTACCAACAATTAGATAT 660
DB 659 GAGCCGGTCAATGCAACGGAGACCCGCTAGTGGCTACTATTCTACCAACAATTAGATAT 718
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTCGAGGTGACAAATTTGACC 720
DB 719 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTCGAGGTGACAAATTTGACC 778
QY 721 TAGCTCCAACTTGAATCAAGATTCAACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
DB 779 TAGCTCCAACTTGAATCAAGATTCAACACAGTTTCTGCTCCAGCTGAATGAGACAATA 838
QY 781 TATACAAGTGGGAAAGAGGCAATACACGGGAAACTAATTTGGAGGTCAACCCCGAA 840
DB 839 TATACAAGTGGGAAAGAGGCAATACACGGGAAACTAATTTGGAGGTCAACCCCGAA 898
```

RESULT 4

```
US-10-491-121-6
; Sequence 6, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 6887
; TYPE: DNA
```

```
QY 841 ATTGATACAAATCGGGAGTGGCCCTTCTGGGAACTAAAAAAACCTCACTAGAAAA 900
DB 899 ATTGATACAAATCGGGAGTGGCCCTTCTGGGAACTAAAAAAACCTCACTAGAAAA 958
QY 901 ATTGCGAGTGAAGATTGTCTTTTACAGTTGTATCAACCGGAGCCAAAAACATCACTGGT 960
DB 959 ATTGCGAGTGAAGATTGTCTTTTACAGTTGTATCAACCGGAGCCAAAAACATCACTGGT 1018
QY 961 CAGAGTCCGGCGCAATCTTCTTCGACCCAGGACCAACAACTGAAGACCACAAA 1020
DB 1019 CAGAGTCCGGCGCAATCTTCTTCGACCCAGGACCAACAACTGAAGACCACAAA 1078
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1080
DB 1079 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAGCT 1138
QY 1081 GCAGTGTGCAATCTAAACAACCTTTGCGCAATCTCCACAGTCCCAATCCCTCAACACC 1140
DB 1139 GCAGTGTGCAATCTAAACAACCTTTGCGCAATCTCCACAGTCCCAATCCCTCAACACC 1198
QY 1141 AAACCCAGTCCGGACCAACAGCACCCCAATAATACCCCGTGTATAAACTTGACATCTCTGAG 1200
DB 1199 AAACCCAGTCCGGACCAACAGCACCCCAATAATACCCCGTGTATAAACTTGACATCTCTGAG 1258
QY 1201 GCAACTCAAGTTGAACAAACATCACCGCAGAAACAGACAAACAGCAGCAGCCTCCGACACT 1260
DB 1259 GCAACTCAAGTTGAACAAACATCACCGCAGAAACAGACAAACAGCAGCAGCCTCCGACACT 1318
QY 1261 CCCTCTGCGACAGCGCAGCGGACCCCAAGAGAGAGAACAGAGAACACCAACAGCAGGAGAGC 1320
DB 1319 CCCTCTGCGACAGCGCAGCGGACCCCAAGAGAGAGAACAGAGAACACCAACAGCAGGAGAGC 1378
QY 1321 ACTGACTTCTGGACCCCGCCACACACAAAGTCCCAAAACCCAGCAGGAGACCCCTGCGC 1380
DB 1379 ACTGACTTCTGGACCCCGCCACACACAAAGTCCCAAAACCCAGCAGGAGACCCCTGCGC 1438
QY 1381 AACAAACAACTCATCAACAAAGATACCGGAGAGAGAGTGCAGCAGCGGAGAGCTAGGC 1440
DB 1439 AACAAACAACTCATCAACAAAGATACCGGAGAGAGAGTGCAGCAGCGGAGAGCTAGGC 1498
QY 1441 TTAATTACCAATATCTATTGCTGGAGTCCAGAGATCTGATCAGAGCGGGAGAGAACTCGA 1500
DB 1499 TTAATTACCAATATCTATTGCTGGAGTCCAGAGCTGATCAGAGCGGGAGAGAACTCGA 1558
QY 1501 CGA 1503
DB 1559 AGA 1561
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-GP(Z) delta GP2 delta C-term A
US-10-491-121-6

Query Match          66.7%; Score 1501.4; DB 20; Length 6887;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGGTACAGGAATATTGAGTTACTCGTGATGATCAAGAGGACATCAATTCCTTT 60
Db 1939 ATGGCGGTACAGGAATATTGAGTTACTCGTGATGATCAAGAGGACATCAATTCCTTT 1998
QY 61 CTTTGGGTAAATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 120
Db 1999 CTTTGGGTAAATTCCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 2058
QY 121 AGCACAATTACAGGTAGTGTGACACAACTAGTTTGTGTCGACAACTGTATCCACA 180
Db 2059 AGCACAATTACAGGTAGTGTGACACAACTAGTTTGTGTCGACAACTGTATCCACA 2118
QY 181 AATCAATTGAGATCAGTTGGACTGTAATCTCGAAGGGAATGAGTGGCACTGACGTGCCA 240
Db 2119 AATCAATTGAGATCAGTTGGACTGTAATCTCGAAGGGAATGAGTGGCACTGACGTGCCA 2178
QY 241 TCTGCAACTTAAAGATGGGGCTTCAAGTCCGCTGTCACCAAGGTGTCATTTATGAA 300
Db 2179 TCTGCAACTTAAAGATGGGGCTTCAAGTCCGCTGTCACCAAGGTGTCATTTATGAA 2238
QY 301 GCTGTGTAATGGGCTGAAACCTGCTACAACTTGAATCAAAAACCTGACGGAGTGCAG 360
Db 2239 GCTGTGTAATGGGCTGAAACCTGCTACAACTTGAATCAAAAACCTGACGGAGTGCAG 2298
QY 361 TGTCTACAGCAGCCGACGAGGATTCGGGCTTCCCGGTGCGGTATGTGCACAAA 420
Db 2299 TGTCTACAGCAGCCGACGAGGATTCGGGCTTCCCGGTGCGGTATGTGCACAAA 2358
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGTGTGCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGCTGAAGGTGTC 2478
QY 541 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 600
Db 2479 GTTGCAATTTCTGATCTGCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 2538
QY 601 GAGCGGTCAATGCAACGGAGGACCGGTCTAGTGCTACTATTTCTACCAATTAGATAT 660
Db 2539 GAGCGGTCAATGCAACGGAGGACCGGTCTAGTGCTACTATTTCTACCAATTAGATAT 2598
QY 661 CAGGCTACCGGTTTGGACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
Db 2599 CAGGCTACCGGTTTGGACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 2658
QY 721 TAGCTCCAACTTGAATCAAGATTCCACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2659 TAGCTCCAACTTGAATCAAGATTCCACCAAGTTTCTGCTCCAGCTGAATGAGACAATA 2718
QY 781 TATACAAGTGGGAAAAGGAGCAATACCAAGGAAAACCTAATTTGGAAGGTCAACCCCGAA 840
Db 2719 TATACAAGTGGGAAAAGGAGCAATACCAAGGAAAACCTAATTTGGAAGGTCAACCCCGAA 2778
QY 841 ATTGATACAACCAATCGGGAGTGGGCTTCTGGGAACTTAAAGAAACCTCACTAGAAA 900
Db 2779 ATTGATACAACCAATCGGGAGTGGGCTTCTGGGAACTTAAAGAAACCTCACTAGAAA 2838
QY 901 ATTTCGAGTGAAGAGTTGTCTTTTCAAGTTGATCAACCGGAGCCAAAACATCAGTGT 960
Db 2839 ATTTCGAGTGAAGAGTTGTCTTTTCAAGTTGATCAACCGGAGCCAAAACATCAGTGT 2898
```

```
QY 961 CAGAGTCCGGCGGAACCTTCTTCGAGCCCGAGGACCAACAACTGAAGACCAAAA 1020
Db 2899 CAGAGTCCGGCGGAACCTTCTTCGAGCCCGAGGACCAACAACTGAAGACCAAAA 2958
QY 1021 ATCATGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1080
Db 2959 ATCATGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 3018
QY 1081 GCAGTGTGGATCTAAACAACCTTTGGCCCAATCTTCCAGAGTCCCAATCCCTCACAACC 1140
Db 3019 GCAGTGTGGATCTAAACAACCTTTGGCCCAATCTTCCAGAGTCCCAATCCCTCACAACC 3078
QY 1141 AAACCAAGTCCGGACACAGACCCATTAATACACCGTGTATTAACCTTGACATCTCTGAG 1200
Db 3079 AAACCAAGTCCGGACACAGACCCATTAATACACCGTGTATTAACCTTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAACTACCGCAGAACAGACAAACGACAGACAGCCTCCGACACT 1260
Db 3139 GCAACTCAAGTTGAACAACTACCGCAGAACAGACAAACGACAGACAGCCTCCGACACT 3198
QY 1261 CCCTCTGCCACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACAGCAAGAGC 1320
Db 3199 CCCTCTGCCACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACAGCAAGAGC 3258
QY 1321 ACTGACTTCTGGACCGGACCAACCAAGTCCCAAAACCAACAGTCCCAAAACCAAGCGCTGGC 1380
Db 3259 ACTGACTTCTGGACCGGACCAACCAAGTCCCAAAACCAACAGTCCCAAAACCAAGCGCTGGC 3318
QY 1381 AACAAACACTCATACCAAGATACCGGAGAGAGTGCAGCAGCGGAGGAGCTAGGC 1440
Db 3319 AACAAACACTCATACCAAGATACCGGAGAGAGTGCAGCAGCGGAGGAGCTAGGC 3378
QY 1441 TTAATTACCAATCTATTTGTGGAGTCCGAGGACTGATCAAGCGGGAGAGAACTCGA 1500
Db 3379 TTAATTACCAATCTATTTGTGGAGTCCGAGGACTGATCAAGCGGGAGAGAACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441

RESULT 5
US-10-491-121-9
; Sequence 9, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ.ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 6914
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-GP(Z) delta TM
US-10-491-121-9

Query Match          66.7%; Score 1501.4; DB 20; Length 6914;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 ATGGCGTTACAGGAATATTGAGTTACCTCTGTGATCGAATCAAGAGGACATCATCTTTT 60
Db 1939 ATGGCGTTACAGGAATATTGAGTTACCTCTGTGATCGAATCAAGAGGACATCATCTTTT 1998
QY 61 CTTTGGGTAAATTTATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAA 120
Db 1999 CTTTGGGTAAATTTATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAA 2058
QY 121 AGCACATTACAGTTAGTGTGACAACTAGTTTGTGCGTACAACTGTCATCTCCACA 180
Db 2059 AGCACATTACAGTTAGTGTGACAACTAGTTTGTGCGTACAACTGTCATCTCCACA 2118
QY 181 AATCAATTAGATCAGTTGGATCTGAACTGTAATCTCGAAGGAATGGAGTGGCACTGACGTGCCA 240
Db 2119 AATCAATTAGATCAGTTGGATCTGAACTGTAATCTCGAAGGAATGGAGTGGCACTGACGTGCCA 2178
QY 241 TCTGCAACTAAAGATGGGGCTTCAAGTCCGGTGTCCACCAAGGTGGTCAATTTATGAA 300
Db 2179 TCTGCAACTAAAGATGGGGCTTCAAGTCCGGTGTCCACCAAGGTGGTCAATTTATGAA 2238
QY 301 GCTGTGTAATGGCTGMAAACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 360
Db 2239 GCTGTGTAATGGCTGMAAACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 2298
QY 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAAA 420
Db 2299 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTGCGGTATGTGCACAAA 2358
QY 421 GTATCAGNAACGGGACCGTGTCCGGAGACTTTGGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGNAACGGGACCGTGTCCGGAGACTTTGGCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGTGAAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTCGTGAAGGTGTC 2478
QY 541 GTTGCAATTTCTGATACGCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 600
Db 2479 GTTGCAATTTCTGATACGCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 2538
QY 601 GAGCGGTCAATGCAACGGGAGACCGGTCTAGTGGCTACTATTCTACCAATTTAGATAT 660
Db 2539 GAGCGGTCAATGCAACGGGAGACCGGTCTAGTGGCTACTATTCTACCAATTTAGATAT 2598
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGCAGAGTACTTCTCGAGTTGACAAATTTGACC 720
Db 2599 CAGGCTACCGGTTTGGAAACCAATGAGCAGAGTACTTCTCGAGTTGACAAATTTGACC 2658
QY 721 TAGTCCAACTTGAATCAAGATTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2659 TAGTCCAACTTGAATCAAGATTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2718
QY 781 TATACAGTGGGAAAGGAGCAATACCACGGGAAACTTAATTTGGAGGTCAACCCCGAA 840
Db 2719 TATACAGTGGGAAAGGAGCAATACCACGGGAAACTTAATTTGGAGGTCAACCCCGAA 2778
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAGAAACCTCACTAGAAAA 900
Db 2779 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAGAAACCTCACTAGAAAA 2838
QY 901 ATTTCAGGTGAAGAGTTGTCTTTTTCAGTTGTATCAAAACGGAGCCCAAAACATCAGTGT 960
Db 2839 ATTTCAGGTGAAGAGTTGTCTTTTTCAGTTGTATCAAAACGGAGCCCAAAACATCAGTGT 2898
QY 961 CAGAGTCCGGCGGAACTTCTTCGGACCCAGGGAACCAACAACTGAAAGCACACAAA 1020
Db 2899 CAGAGTCCGGCGGAACTTCTTCGGACCCAGGGAACCAACAACTGAAAGCACACAAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTTCTCTGCAATGGTTTCAAGTGCACAGTCAAGGAAGGGAAGCT 1080
Db 2959 ATCATGGCTTCAGAAAAATTTCTCTGCAATGGTTTCAAGTGCACAGTCAAGGAAGGGAAGCT 3018
QY 1081 GCAGTGTGCAATCTAAACAAACCTTGGCCAAATCTCCACAGGTCCCCCAATCCCTCAACACC 1140

Db 3019 GCAGTGTGCAATCTAAACAAACCTTTGCCACAATCTCCACGAGTCCCAATCCCTCAACACC 3078
QY 1141 AAACAGGTCCGGACAAACAGCACCAATAATACACCGTGTATAAACTTGACATCTCTGAG 1200
Db 3079 AAACAGGTCCGGACAAACAGCACCAATAATACACCGTGTATAAACTTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAACTATCCCGCAGAAACAGACAAACAGACAGCCTCGACACT 1260
Db 3139 GCAACTCAAGTTGAACAACTATCCCGCAGAAACAGCACCAACAGCAGCCTCGACACT 3198
QY 1261 CCCTCTGCACAGCGCAGCCGGACCCCAAAAGCAGAGAACAACCAACAGCAGCAGAGC 1320
Db 3199 CCCTCTGCACAGCGCAGCCGGACCCCAAAAGCAGAGAACAACCAACAGCAGCAGAGC 3258
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAACTGTCCTCCCAAAACCAACAGCAGCAGCCTCGC 1380
Db 3259 ACTGACTTCTGGACCCCGCCACCAACAACTGTCCTCCCAAAACCAACAGCAGCAGCCTCGC 3318
QY 1381 AACAACTACTATCTACCAAGATACCGGAGAGAGTGCAGCAGCGGGAAGCTAGGC 1440
Db 3319 AACAACTACTATCTACCAAGATACCGGAGAGAGTGCAGCAGCGGGAAGCTAGGC 3378
QY 1441 TTAATTACCAATATCTATGCTGGAGTCCAGACTGATCAGCGGGGAGAGACTCGA 1500
Db 3379 TTAATTACCAATATCTATGCTGGAGTCCAGACTGATCAGCGGGGAGAGACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441

RESULT 6

US-10-491-121-7
; Sequence 7, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 7044
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-Gp(Z) delta GP2 Delta C-term B
US-10-491-121-7

Query Match 66.7%; Score 1501.4; DB 20; Length 7044;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGCGTTACAGGAATATTGAGTTACCTCTGTGATCGAATCAAGAGGACATCATCTTTT 60
Db 1939 ATGGCGTTACAGGAATATTGAGTTACCTCTGTGATCGAATCAAGAGGACATCATCTTTT 1998
QY 61 CTTTGGGTAAATTTATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCACAA 120
Db 1999 CTTTGGGTAAATTTATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCAA 2058
QY 121 AGCACATTACAGTTAGTGTGACAAACTAGTTTGTGCTGAGCAAACTGTCATCCACA 180

```
Db 2059 AGCACTTACAGGTAGTGTGATGTCACAAACTAGTGTGTCGACAAACTGTCTATCCACA 2118
QY 181 AATCAATTGAGATCAGTTGGACTGTAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db 2119 AATCAATTGAGATCAGTTGGACTGTAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 2178
QY 241 TCTGCACTAATAAGATGGGGCTTCAGGTCGGGTGTCACCAAAAGGTGGTCAATTATGAA 300
Db 2179 TCTGCACTAATAAGATGGGGCTTCAGGTCGGGTGTCACCAAAAGGTGGTCAATTATGAA 2238
QY 301 GCTGTGTAATGGGCTGAAAACCTGCTACAACTCTTGAATCAAAACCTGACGGAGTGAG 360
Db 2239 GCTGTGTAATGGGCTGAAAACCTGCTACAACTCTTGAATCAAAACCTGACGGAGTGAG 2298
QY 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTCGGGTATGTGCACAAA 420
Db 2299 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTCGGGTATGTGCACAAA 2358
QY 421 GTATCAGGAAACGGGACCGGTGTCGGGAGACTTTGGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAAACGGGACCGGTGTCGGGAGACTTTGGCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGGACGACTTTCGCTGAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTATCTACCGAGGAGCAGTTCGCTGAGGTGTC 2478
QY 541 GTTGCAATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 600
Db 2479 GTTGCAATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAGCTCACACCCCTTGAGA 2538
QY 601 GAGCCGGTCAATGCAACGGAGGACCGGTCTAGTGCTACTATTTACCACAATTAGATAT 660
Db 2539 GAGCCGGTCAATGCAACGGAGGACCGGTCTAGTGCTACTATTTACCACAATTAGATAT 2598
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGTTGCAATTTGACC 720
Db 2599 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGTTGCAATTTGACC 2658
QY 721 TACGTCCTCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2659 TACGTCCTCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2718
QY 781 TATACAAGTGGGAAAGAGGCAATACCAACGGGAAACCTAAATTTGGAAGGTCAACCCCGAA 840
Db 2719 TATACAAGTGGGAAAGAGGCAATACCAACGGGAAACCTAAATTTGGAAGGTCAACCCCGAA 2778
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAAGAACTTAAAGAACTCACTAGAAA 900
Db 2779 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAAGAACTTAAAGAACTCACTAGAAA 2838
QY 901 ATTCGAGTGAAGAGTTGTCTTTTACAGTTGTATCAAAACGGAGCCAAAACATCAGTGGT 960
Db 2839 ATTCGAGTGAAGAGTTGTCTTTTACAGTTGTATCAAAACGGAGCCAAAACATCAGTGGT 2898
QY 961 CAGAGTCCGGCCGGAACCTTTCTTCGACCCAGGACCAACACAACTTGAAGACCAAAA 1020
Db 2899 CAGAGTCCGGCCGGAACCTTTCTTCGACCCAGGACCAACACAACTTGAAGACCAAAA 2958
QY 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAGCT 1080
Db 2959 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAGCT 3018
QY 1081 GCAGTGTGCACTCTAACAACTTTCGCAAACTCTCACAGGTCGCAATTCCTCTCAACACC 1140
Db 3019 GCAGTGTGCACTCTAACAACTTTCGCAAACTCTCACAGGTCGCAATTCCTCTCAACACC 3078
QY 1141 AAACCAAGTCCGGACAAACAGCAGCCCAATATACACCCGCTGTATATACTTGAATCTCTGAG 1200
Db 3079 AAACCAAGTCCGGACAAACAGCAGCCCAATATACACCCGCTGTATATACTTGAATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAACTACCGCAGACAGACCAAGGAGGAGTCCGACACT 1260
```

```
Db 3139 GCAACTCAAGTTGAAACAATCATCCGCGAGAACACAGCAACACAGCAGCCTCCGACACT 3198
QY 1261 CCCTCTGCCACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACACGAGCAAGAGC 1320
Db 3199 CCCTCTGCCACGACCGCAGCGGACCCCAAAAGCAGAGAACACCAACACGAGCAAGAGC 3258
QY 1321 ACTGACTTCTGGACCCCGCCACCAACAAGTCCCAAAACCAAGGAGACCGCTGGC 1380
Db 3259 ACTGACTTCTGGACCCCGCCACCAACAAGTCCCAAAACCAAGGAGACCGCTGGC 3318
QY 1381 AACAAACAACACTCATCAACCAAGTACCGGAGAGAGAGTGCAGCAGCGGGAAGCTAGGC 1440
Db 3319 AACAAACAACACTCATCAACCAAGTACCGGAGAGAGAGTGCAGCAGCGGGAAGCTAGGC 3378
QY 1441 TTAATTACCAATACTATTGTCTGGAGTCCGAGGACTGATCAAGCGGAGAGAACTCGA 1500
Db 3379 TTAATTACCAATACTATTGTCTGGAGTCCGAGGACTGATCAAGCGGAGAGAACTCGA 3438
QY 1501 CGA 1503
Db 3439 AGA 3441

RESULT 7
US-10-491-121-8
; Sequence 8, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 7106
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-GP(2) delta GP2 delta FUS
US-10-491-121-8

Query Match 66.7%; Score 1501.4; DB 20; Length 7106;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGCGTTACAGGAATATTTCAGTTACTCTCGTGATGATTCAAGAGAGACATCATCTTTT 60
Db 1939 ATGGGCGTTACAGGAATATTTCAGTTACTCTCGTGATGATTCAAGAGAGACATCATCTTTT 1998
QY 61 CTTTGGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCAAAT 120
Db 1999 CTTTGGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCAATCCAAAT 2058
QY 121 AGCACTTACAGGTAGTGTGATGTCGACAAACTAGTGTGTCGACAAACTGTCTATCCACA 180
Db 2059 AGCACTTACAGGTAGTGTGATGTCGACAAACTAGTGTGTCGACAAACTGTCTATCCACA 2118
QY 181 AATCAATTGAGATCAGTTGGAGTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db 2119 AATCAATTGAGATCAGTTGGAGTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 2178
QY 241 TCTGCACTAATAAGATGGGGCTTTCAGGTCGGGTGTCACCAAGGAGGAGTCCGACACT 300
```

```
Db 2179 TCTGCAACTAAAGATGGGGCTTCAGGTCGGGTGTCCTCCCAACCAAGGGTGTCAATTATGAA 2238
QY 301 GCTGTGTAATGGGCTGAAACTGCTACAACTTTGAAATCAAAAACCTGACGGGAGTGAG 360
Db 2239 GCTGTGTAATGGGCTGAAACTGCTACAACTTTGAAATCAAAAACCTGACGGGAGTGAG 2298
QY 361 TGTCTACAGCAGCGCAGACGGGATTCGGGGCTTCCCGCGTGCCGGTATGTGCACAAA 420
Db 2299 TGTCTACAGCAGCGCAGACGGGATTCGGGGCTTCCCGCGTGCCGGTATGTGCACAAA 2358
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 2359 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 2418
QY 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGTGAAGGTGTC 540
Db 2419 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGTGAAGGTGTC 2478
QY 541 GTTGCATTTCTGATACTGCCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 600
Db 2479 GTTGCATTTCTGATACTGCCCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 2538
QY 601 GAGCCGGTCAATGCAACCGGAGGACCGCTCTAGTGGCTACTATTCTACCAANTTAGATAT 660
Db 2539 GAGCCGGTCAATGCAACCGGAGGACCGCTCTAGTGGCTACTATTCTACCAANTTAGATAT 2598
QY 661 CAGGCTACCGGTTTTGGAACCAATGAGACAGAGTACTTCTTCGAGGTTGACAAATTGACC 720
Db 2599 CAGGCTACCGGTTTTGGAACCAATGAGACAGAGTACTTCTTCGAGGTTGACAAATTGACC 2658
QY 721 TAGCTCAACTTGAATCAAGATTTCACCAAGTTTCTGCTCTCAGCTGAATGAGACAATA 780
Db 2659 TAGCTCAACTTGAATCAAGATTTCACCAAGTTTCTGCTCTCAGCTGAATGAGACAATA 2718
QY 781 TATCAAGTGGGAAAGAGGAGCAATACCGGGAACAACTTAATTTGGAAGTCAACCCGGA 840
Db 2719 TATCAAGTGGGAAAGAGGAGCAATACCGGGAACAACTTAATTTGGAAGTCAACCCGGA 2778
QY 841 ATTGATACAACAATCGGGAGTGGGCTTCTCGGAAACTTAAAGAAACCTCACTAGAAAA 900
Db 2779 ATTGATACAACAATCGGGAGTGGGCTTCTCGGAAACTTAAAGAAACCTCACTAGAAAA 2838
QY 901 ATTGCGAGTGAAGATGTCTTTTCAAGTTGTATCAACCGGAGCCAAAAACATCAGTGT 960
Db 2839 ATTGCGAGTGAAGATGTCTTTTCAAGTTGTATCAACCGGAGCCAAAAACATCAGTGT 2898
QY 961 CAGAGTCCGGCGGGAATCTTCCGACCCAGGACCAACACAACTGAAGACACACAA 1020
Db 2899 CAGAGTCCGGCGGGAATCTTCCGACCCAGGACCAACACAACTGAAGACACACAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTCCTGTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 1080
Db 2959 ATCATGGCTTCAGAAAAATTCCTGTGCAATGGTTCAAGTGCACAGTCAAGGAAGGAAAGCT 3018
QY 1081 GCAGTGTGCGATCTAAACAACCTTTGCAAAATCTCCACGAGTCCCAATCCCTCAACAC 1140
Db 3019 GCAGTGTGCGATCTAAACAACCTTTGCAAAATCTCCACGAGTCCCAATCCCTCAACAC 3078
QY 1141 AAACCAAGTCCGGACAGACGCCATTAATACCCGCTGTATAAATCTGACATCTCTGAG 1200
Db 3079 AAACCAAGTCCGGACAGACGCCATTAATACCCGCTGTATAAATCTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAACATACCCGAGAACAGACAGACGACAGCCCTCCGACACT 1260
Db 3139 GCAACTCAAGTTGAACAACATACCCGAGAACAGACAGACGACGACGCTCCGACACT 3198
QY 1261 CCTCTGCCACGACCGGACCGGACCCCAAGAGCAGAGAAACCAACACGAGCAGAGAGC 1320
Db 3199 CCTCTGCCACGACCGGACCGGACCCCAAGAGCAGAGAAACCAACACGAGCAGAGAGC 3258
QY 1321 ACTGACTTCTTGACCCCGGACCCCAAGAGTCCCAAAACCAACGAGCAGAGCCGCTGCG 1380
Db 3259 ACTGACTTCTTGACCCCGGACCCCAAGAGTCCCAAAACCAACGAGCAGAGCCGCTGCG 3318
```

RESULT 8

```
US-10-491-121-1
; Sequence 1, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 7154
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012 -GP(Z)
US-10-491-121-1
```

```
Query Match 66.7%; Score 1501.4; DB 20; Length 7154;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGGTTACAGGAATATTCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 60
Db 1939 ATGGCGGTTACAGGAATATTCAGTTACCTCGTGATCGATTCAAGAGGACATCATTTCTTT 1998
QY 61 CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCATCTTGGAGTCAATCCACAAT 120
Db 1999 CTTTGGGTAATTATCTTTTCCAAAGAACATTTTCCATCCCATCTTGGAGTCAATCCACAAT 2058
QY 121 AGCATTACAGGTTAGTGATGTGCAAAAATAGTTTGTGTGCAAAAATGTCTCAATCCACA 180
Db 2059 AGCATTACAGGTTAGTGATGTGCAAAAATAGTTTGTGTGCAAAAATGTCTCAATCCACA 2118
QY 181 AATCAATTTGAGATCAGTTGGACTGAACTCTCGAAGGGAATGGAAGTGGCAACTGACGTGCA 240
Db 2119 AATCAATTTGAGATCAGTTGGACTGAACTCTCGAAGGGAATGGAAGTGGCAACTGACGTGCA 2178
QY 241 TCTGCAACTAAAGATGGGGCTTTCAGGTCGGGTGTCCTCCCAAGGTTGGTCAATTATGAA 300
Db 2179 TCTGCAACTAAAGATGGGGCTTTCAGGTCGGGTGTCCTCCCAAGGTTGGTCAATTATGAA 2238
QY 301 GCTGTGTAATGGGCTGAAAACTGCTTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 360
Db 2239 GCTGTGTAATGGGCTGAAAACTGCTTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 2298
QY 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGCCGGTATGTGCACAAA 420
Db 2299 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGCGTGCCGGTATGTGCACAAA 2358
```



```
QY 541 GTTGCATTCTTGATCTATGCCCCAGCTAAGAGGAGCTTCTTCAGCTCAACCCCTTGAGA 600
Db 2479 GTTGCATTCTTGATCTATGCCCCAGCTAAGAGGAGCTTCTTCAGCTCAACCCCTTGAGA 2538
QY 601 GAGCCGGTCAATGCAACGAGGAGCGCTCTAGTGGCTACTATTCTACCAATATTAGATAT 660
Db 2539 GAGCCGGTCAATGCAACGAGGAGCGCTCTAGTGGCTACTATTCTACCAATATTAGATAT 2598
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTTCGAGGTGGACAAATTTGACC 720
Db 2599 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTCTTCGAGGTGGACAAATTTGACC 2658
QY 721 TAGCTCCAACTTGAATCAAGATTCAACACAGTCTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2659 TAGCTCCAACTTGAATCAAGATTCAACACAGTCTTCTGCTCCAGCTGAATGAGACAATA 2718
QY 781 TATACAAGTGGGAAGGAGCAATACACCGGGAACCTAATTTGGAAGTCAACCCCGAA 840
Db 2719 TATACAAGTGGGAAGGAGCAATACACCGGGAACCTAATTTGGAAGTCAACCCCGAA 2778
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAAAAAACCTCACTAGAAAA 900
Db 2779 ATTGATACAACTCGGGAGTGGGCTTCTGGGAACTTAAAAAAACCTCACTAGAAAA 2838
QY 901 ATTCGAGTGAAGAGTGTCTTTTCAAGTTGTATCAACGGAGCCAAAAACATCATGTGT 960
Db 2839 ATTCGAGTGAAGAGTGTCTTTTCAAGTTGTATCAACGGAGCCAAAAACATCATGTGT 2898
QY 961 CAGAGTCCGGCGGAGTCTTCCGACCCAGGAGCCCAACACAACTGAAGACACACAA 1020
Db 2899 CAGAGTCCGGCGGAGTCTTCCGACCCAGGAGCCCAACACAACTGAAGACACACAA 2958
QY 1021 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGAAAGGGAAGCT 1080
Db 2959 ATCATGGCTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGAAAGGGAAGCT 3018
QY 1081 GCAGTGTCCGATCTAAACACCTTTGCGCAAACTCTCCAGAGTCCCAATCCCTCAACAC 1140
Db 3019 GCAGTGTCCGATCTAAACACCTTTGCGCAAACTCTCCAGAGTCCCAATCCCTCAACAC 3078
QY 1141 AAACGAGTCCGGACACAGCAGCCCAATATACCCGTTGTATAAATTTGACATCTCTGAG 1200
Db 3079 AAACGAGTCCGGACACAGCAGCCCAATATACCCGTTGTATAAATTTGACATCTCTGAG 3138
QY 1201 GCAACTCAAGTTGAACAAATCAACGAGAGACAGACAGCAGCAGCAGCTCCGACACT 1260
Db 3139 GCAACTCAAGTTGAACAAATCAACGAGAGACAGACAGCAGCAGCAGCTCCGACACT 3198
QY 1261 CCTCTGCCACGACCGGAGCGGAGACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 3199 CCTCTGCCACGACCGGAGCGGAGACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3258
QY 1321 ACTGATCTCTGGACCCCGGACCAACAACTCCCAAAACAGAGAGAGAGAGAGAGAGAG 1380
Db 3259 ACTGATCTCTGGACCCCGGACCAACAACTCCCAAAACAGAGAGAGAGAGAGAGAGAG 3318
QY 1381 AACCAACACATCATCAACAGATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 3319 AACCAACACATCATCAACAGATACCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3378
QY 1441 TTAATTTACCAATACTATTCTGAGTCCGAGAGTGTATCAAGGGGGGAGAGAGAGAGAG 1500
Db 3379 TTAATTTACCAATACTATTCTGAGTCCGAGAGTGTATCAAGGGGGGAGAGAGAGAGAG 3438
QY 1501 CGA 1503
Db 3439 AGA 3441
```

RESULT 10
US-10-860-878-3
; Sequence 3, Application US/10860878
; Publication No. US20050130129A1

```
; GENERAL INFORMATION:  
; APPLICANT: Nabel, Gary J  
; APPLICANT: Delgado, Rafael  
; APPLICANT: Yang, Zhi-yong  
; TITLE OF INVENTION: Targeting Gene Transfer Vectors To Certain Cell Types  
; TITLE OF INVENTION: By Pseudotyping With Viral Glycoprotein  
; FILE REFERENCE: 21158-001474FOA  
; CURRENT APPLICATION NUMBER: US/10/860,878  
; PRIOR FILING DATE: 2004-06-03  
; PRIOR APPLICATION NUMBER: US/09/600,766  
; PRIOR FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: USSN 60/072033  
; PRIOR FILING DATE: 1998-01-21  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 7285  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:Plasmid  
; OTHER INFORMATION: containing DNA of GP of Ebola Virus, Zaire strain  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (886)..(1129)  
; OTHER INFORMATION: CMV IE 5' UT  
; FEATURE:  
; NAME/KEY: intron  
; LOCATION: (1130)..(1840)  
; OTHER INFORMATION: CMV IE INT  
; FEATURE:  
; NAME/KEY: enhancer  
; LOCATION: (248)..(885)  
; OTHER INFORMATION: CMV enhancer  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (1870)..(4301)  
; OTHER INFORMATION: GP (Z)  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (4302)..(4854)  
; OTHER INFORMATION: TGGH  
; FEATURE:  
; NAME/KEY: gene  
; LOCATION: (6350)..(6972)  
; OTHER INFORMATION: Kan r  
; US-10-860-878-3  
  
Query Match 66.7%; Score 1501.4; DB 22; Length 7285;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ATGGGCGTTACAGGAATATTCAGTTACCTCGTGATCCGATTCAGAGGACATCATTTCTTT 60  
Db 2072 ATGGGCGTTACAGGAATATTCAGTTACCTCGTGATCCGATTCAGAGGACATCATTTCTTT 2131  
QY 61 CTTTGGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 120  
Db 2132 CTTTGGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCACTTGGAGTCATCCACAAT 2191  
QY 131 AGCAATTTACAGGTAGTGTGACAACTAGTTTCTCGTGACAAACTGTGTCAATCCACA 180  
Db 2192 AGCAATTTACAGGTAGTGTGACAACTAGTTTCTCGTGACAAACTGTGTCAATCCACA 2251  
QY 181 AATCAATTTGAGATCAGTTGGATCTCCAGGGAATGGAGTGGCACTGACCTGCCA 240  
Db 2252 AATCAATTTGAGATCAGTTGGATCTCCAGGGAATGGAGTGGCACTGACCTGCCA 2311  
QY 241 TCTGCAACTAAAGATGGGGCTTCAGGTCCGCTGCCCAACCAAGGTGGTCAATTTATGAA 300  
Db 2312 TCTGCAACTAAAGATGGGGCTTCAGGTCCGCTGCCCAACCAAGGTGGTCAATTTATGAA 2371  
QY 301 GCTGGTGAATGGGCTGGAANAATGCTGTACAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
```



```
Db 1756 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATAAAGAGGGTGTCTTCTTC 1815
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
Db 1816 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 1875
Qy 541 GTTGCATTTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCAAGCTCACAACCCCTTGAGA 600
Db 1876 GTTGCATTTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCAAGCTCACAACCCCTTGAGA 1935
Qy 601 GAGCCGGTCAATGCAACCGGAGACCGGTCTAGTGTCTACTATTCTACCAATATAGATAT 660
Db 1936 GAGCCGGTCAATGCAACCGGAGACCGGTCTAGTGTCTACTATTCTACCAATATAGATAT 1995
Qy 661 CAGGCTACCGGTTTGGGAACCAATGAGACAGAGTACTTGTTCGAGGTGGACAAATTTGACC 720
Db 1996 CAGGCTACCGGTTTGGGAACCAATGAGACAGAGTACTTGTTCGAGGTGGACAAATTTGACC 2055
Qy 721 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2056 TAGCTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2115
Qy 781 TATACAAGTGGGAAAGAGGAGCAATACCAAGGAAACTTAATTTGGAAGGTCAACCCGAA 840
Db 2116 TATACAAGTGGGAAAGAGGAGCAATACCAAGGAAACTTAATTTGGAAGGTCAACCCGAA 2175
Qy 841 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAGAACTCACTAGAGAAA 900
Db 2176 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAGAACTCACTAGAGAAA 2235
Qy 901 ATTCCAGTGAAGAGTGTCTTTTCAAGTTGATCAACGGAGCCAAACCAATCATAGTGGT 960
Db 2236 ATTCCAGTGAAGAGTGTCTTTTCAAGTTGATCAACGGAGCCAAACCAATCATAGTGGT 2295
Qy 961 CAGAGTCCGGCCGGAATCTTCCGACCCAGGAGCAACACAACTGAAGACACACAA 1020
Db 2296 CAGAGTCCGGCCGGAATCTTCCGACCCAGGAGCAACACAACTGAAGACACACAA 2355
Qy 1021 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 1080
Db 2356 ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGGAAGCT 2415
Qy 1081 GCAGTGTGCACTATACAAACCCCTTGCCACAATCTCCACGAGTCCCAATCCCTCAACAC 1140
Db 2416 GCAGTGTGCACTATACAAACCCCTTGCCACAATCTCCACGAGTCCCAATCCCTCAACAC 2475
Qy 1141 AAACAGGTCCGGACACAGCACCCTAATACACCGGTGTATAAATTTGACATCTCTGAG 1200
Db 2476 AAACAGGTCCGGACACAGCACCCTAATACACCGGTGTATAAATTTGACATCTCTGAG 2535
Qy 1201 GCAACTCAGTGTGAACAACTACCGCAGAACAGACAGCAGCAGCCTCCGACACT 1260
Db 2536 GCAACTCAGTGTGAACAACTACCGCAGAACAGACAGCAGCAGCCTCCGACACT 2595
Qy 1261 CCCTCTGCCACGACCGCAGCCGACCCCAAGAGAGAGAACCAACACGAGCAGGAGAGC 1320
Db 2596 CCCTCTGCCACGACCGCAGCCGACCCCAAGAGAGAGAACCAACACGAGCAGGAGAGC 2655
Qy 1321 ACTGACTTCTGGACCCCGCACCACCAACAGTCCCAAAACCAACGCGAGACCGCTGGC 1380
Db 2656 ACTGACTTCTGGACCCCGCACCACCAACAGTCCCAAAACCAACGCGAGACCGCTGGC 2715
Qy 1381 AACAAACACACTCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGGAGCTAGGC 1440
Db 2716 AACAAACACACTCATCAACAGATACCGGAGAGAGAGTCCAGCAGCGGGAGCTAGGC 2775
Qy 1441 TTAATTACCAATATCTATTCTGGAGTCCGAGGACTGATCAGAGCGGGAGAGAACTCGA 1500
Db 2776 TTAATTACCAATATCTATTCTGGAGTCCGAGGACTGATCAGAGCGGGAGAGAACTCGA 2835
Qy 1501 CGA 1503
Db 2836 AGA 2838
```

RESULT 12

```
US-10-491-121-28
; Sequence 28, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; FILE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 8439
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pAdapt Ebola GP(Z)
US-10-491-121-28
```

```
Query Match 66.7%; Score 1501.4; DB 20; Length 8439;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGGGCGTTACAGGAATATTCGACTTACCTCGTGATCGATTCCATTCAGAGGACATCATCTTT 60
Db 1336 ATGGGCGTTACAGGAATATTCGACTTACCTCGTGATCGATTCCATTCAGAGGACATCATCTTT 1395
Qy 61 CTTTGGGTAATATTCCTTTTCCAAAGAAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 120
Db 1396 CTTTGGGTAATATTCCTTTTCCAAAGAAACATTTTCCATCCCACTTGGAGTCAATCCACAAT 1455
Qy 121 AGCACTTACAGGTAGTGTGACAAATAGTGTTCGTGACAAATCTGTATCCACA 180
Db 1456 AGCACTTACAGGTAGTGTGACAAATAGTGTTCGTGACAAATCTGTATCCACA 1515
Qy 181 AATCAATTTGAGATCAGTTGGACTGATCTCCAGGGAATGGAGTGGCACTGACCTGCCA 240
Db 1516 AATCAATTTGAGATCAGTTGGACTGATCTCCAGGGAATGGAGTGGCACTGACCTGCCA 1575
Qy 241 TCTGCAACTAAAGATGGGCTTCAGGTCCGGTGTCCCAACCAAGGTGGTCAATTTATGAA 300
Db 1576 TCTGCACTAAAGATGGGCTTCAGGTCCGGTGTCCCAACCAAGGTGGTCAATTTATGAA 1635
Qy 301 GCTGTGAATGGGCTGAAAACTGTGTAATCTTGAATCAAAAAACCTGACGGGAGTGAG 360
Db 1636 GCTGTGAATGGGCTGAAAACTGTGTAATCTTGAATCAAAAAACCTGACGGGAGTGAG 1695
Qy 361 TGTCTACAGCAGCGCCAGCGGATTCGGGCTTCGCCCGGTTCGCCGTGTGTGTGACAAA 420
Db 1696 TGTCTACAGCAGCGCCAGCGGATTCGGGCTTCGCCCGGTTCGCCGTGTGTGTGACAAA 1755
Qy 421 GTATCAGGAACGGGACCGGTGTGCCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 1756 GTATCAGGAACGGGACCGGTGTGCCGGAGACTTTGCTTCCATAAAGAGGGTCTTTCTTC 1815
Qy 481 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 540
Db 1816 CTGTATGATCGACTTGTCTTCCACAGTTATCTACCGAGGAACGACTTTTCGCTGAAGGTGTC 1875
Qy 541 GTTGCATTTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCAAGCTCACAACCCCTTGAGA 600
Db 1876 GTTGCATTTCTGATATCTGCCCCAAGCTAAGAGGACTTCTTCAAGCTCACAACCCCTTGAGA 1935
```

```
QY 601 GAGCGGTCAATGCAACGAGGACCGCTCTAGTGGCTACTATTCTTACCACAAATTAGATAT 660
Db 1936 GAGCGCGTCAATGCAACGAGGACCGCTCTAGTGGCTACTATTCTTACCACAAATTAGATAT 1995
QY 661 CAGGCTACCGGTTTGGAAACCAATCAGACAGAGTACTTGTTCGAGGTTTGACAATTGACC 720
Db 1996 CAGGCTACCGGTTTGGAAACCAATCAGACAGAGTACTTGTTCGAGGTTTGACAATTGACC 2055
QY 721 TACGTCCTCAATGGAATCAAGATTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
Db 2056 TACGTCCTCAATGGAATCAAGATTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2115
QY 781 TATACAAGTGGGAAAGAGGCAATACACGCGGAAACCTAATTTGGAAGGTCAACCCCGAA 840
Db 2116 TATACAAGTGGGAAAGAGGCAATACACGCGGAAACCTAATTTGGAAGGTCAACCCCGAA 2175
QY 841 ATTGATACAACTCGGAGGTGGGCTTCTCGGAAACCTTAAAGAACTCACTAGAAAA 900
Db 2176 ATTGATACAACTCGGAGGTGGGCTTCTCGGAAACCTTAAAGAACTCACTAGAAAA 2235
QY 901 ATTGAGTGAAGAGTGTCTTTTACAGTTGTATCAAAACGGACCAAAACATCAGTGGT 960
Db 2236 ATTGAGTGAAGAGTGTCTTTTACAGTTGTATCAAAACGGACCAAAACATCAGTGGT 2295
QY 961 CAGAGTCCGGCGGAACTTCTTCGACCCAGGGACCAACAACTGAAGACCAACAA 1020
Db 2296 CAGAGTCCGGCGGAACTTCTTCGACCCAGGGACCAACAACTGAAGACCAACAA 2355
QY 1021 ATCATGGTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGAACT 1080
Db 2356 ATCATGGTTCAGAAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAGGAACT 2415
QY 1081 GCAGTGTGCGACTTAACAAACCTTGCACAACTCTCCAGAGTCCCAATCCCTCAACACC 1140
Db 2416 GCAGTGTGCGACTTAACAAACCTTGCACAACTCTCCAGAGTCCCAATCCCTCAACACC 2475
QY 1141 AAACAGGTCCGGACAAACAGCAACCCATTAACACCCGTTGATATAAATTGACATCTCTGAG 1200
Db 2476 AAACAGGTCCGGACAAACAGCAACCCATTAACACCCGTTGATATAAATTGACATCTCTGAG 2535
QY 1201 GCAACTCAAGTTGAACATCAACCGCAGACAGACAGACAGACAGACAGACAGTCCGACACT 1260
Db 2536 GCAACTCAAGTTGAACATCAACCGCAGACAGACAGACAGACAGACAGACAGTCCGACACT 2595
QY 1261 CCCTCTGCGACAGCGCAGCGGACCCCAAGCAGAGAACACCAACAGCAGCAGCAGCAGTGGC 1320
Db 2596 CCCTCTGCGACAGCGCAGCGGACCCCAAGCAGAGAACACCAACAGCAGCAGCAGCAGTGGC 2655
QY 1321 ACTGACTTCTGGAACCCCGCACCAACAAAGTCCCAAAACCAACAGCAGCAGCAGCAGTGGC 1380
Db 2656 ACTGACTTCTGGAACCCCGCACCAACAAAGTCCCAAAACCAACAGCAGCAGCAGCAGTGGC 2715
QY 1381 AACAAACACTCATCAACAGATACCGGAGAGAGAGTCCGAGCAGCAGCAGCAGTGGC 1440
Db 2716 AACAAACACTCATCAACAGATACCGGAGAGAGAGTCCGAGCAGCAGCAGCAGTGGC 2775
QY 1441 TTAATTACCAATATATTGCTGGAGTCCGAGGACTGATCAGAGGCGGAGAGAACTCGA 1500
Db 2776 TTAATTACCAATATATTGCTGGAGTCCGAGGACTGATCAGAGGCGGAGAGAACTCGA 2835
QY 1501 CGA 1503
Db 2836 AGA 2838
```

RESULT 13

```
US-10-491-121-25
; Sequence 25, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
```

```
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 10783
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct PAD/CMV-GP(dTM) (Z-CITE-S)
US-10-491-121-25
```

Query Match 66.7%; Score 1501.4; DB 20; Length 10783;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGGGCGTTACAGGAATATTCAGTTACCTCGTGATCGATTCAAGAGGACATCATCTTTT 60
Db 1324 ATGGGCGTTACAGGAATATTCAGTTACCTCGTGATCGATTCAAGAGGACATCATCTTTT 1383
QY 61 CTTTGGGTAAATATCTCTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACAAAT 120
Db 1384 CTTTGGGTAAATATCTCTTTCCAAAGAACATTTTCCATCCCACTTGGAGTATCCACAAAT 1443
QY 121 AGCATTACAGGTTAGTGATGTCGACAAACTAGTTTGTGTCGACAAACTGTCATCCACA 180
Db 1444 AGCATTACAGGTTAGTGATGTCGACAAACTAGTTTGTGTCGACAAACTGTCATCCACA 1503
QY 181 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
Db 1504 AATCAATTGAGATCAGTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 1563
QY 241 TCTGCNACTAAAGATGGGGCTTCAGGTCGGGTGCCCAAGGTTGGTCAATATATGAA 300
Db 1564 TCTGCNACTAAAGATGGGGCTTCAGGTCGGGTGCCCAAGGTTGGTCAATATATGAA 1623
QY 301 GCTGTGAATCGGCTGAAAACCTGCTACAACTCTGAAATCAAAAACCTGACGGGAGTGAG 360
Db 1624 GCTGTGAATCGGCTGAAAACCTGCTACAACTCTGAAATCAAAAACCTGACGGGAGTGAG 1683
QY 361 TGTCTACAGACGCGCCAGACGGGATTCGGGGCTTCCCGGTCGCGGTATGTGCACAAA 420
Db 1684 TGTCTACAGACGCGCCAGACGGGATTCGGGGCTTCCCGGTCGCGGTATGTGCACAAA 1743
QY 421 GTATCAGGAACGGGACCGTGTGCGGAGACTTTCGCTTCCATAAAGAGGGTCTTTCTTC 480
Db 1744 GTATCAGGAACGGGACCGTGTGCGGAGACTTTCGCTTCCATAAAGAGGGTCTTTCTTC 1803
QY 481 CTGTATGATCGACTTGTCTTCACAGTTATCTACAGGAAACGACTTTCGCTGAAGGTGTC 540
Db 1804 CTGTATGATCGACTTGTCTTCACAGTTATCTACAGGAAACGACTTTCGCTGAAGGTGTC 1863
QY 541 GTTGCATTTCTGTACTGCTCCCAAGCTAAAGAGGACTTCTTTTTCAGCTCAACCCCTTGAGA 600
Db 1864 GTTGCATTTCTGTACTGCTCCCAAGCTAAAGAGGACTTCTTTTTCAGCTCAACCCCTTGAGA 1923
QY 601 GAGCGGTCAATGCAACGAGGACCCGCTCTAGTGGCTACTATTCTTACACAAATTAGATAT 660
Db 1924 GAGCGGTCAATGCAACGAGGACCCGCTCTAGTGGCTACTATTCTTACACAAATTAGATAT 1983
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGACTTGTTCGAGGTTGACAATTGACC 720
Db 1984 CAGGCTACCGGTTTGGAAACCAATGAGACAGACTTGTTCGAGGTTGACAATTGACC 2043
```

```
QY 721 TACGTCACAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
DB |||||
QY 2044 TACGTCACAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 2103
DB |||||
QY 781 TATACAGTGGGAAAGGAGCAATACCAGGGGAAACTTAATTTGGAAGGTCAACCCCGAA 840
DB |||||
QY 2104 TATACAGTGGGAAAGGAGCAATACCAGGGGAAACTTAATTTGGAAGGTCAACCCCGAA 2163
DB |||||
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTCGGAAACTTAAAGGAACTTAAAGGAACTTAAAGGAA 900
DB |||||
QY 2164 ATTGATACAACTCGGGAGTGGGCTTCTCGGAACTTAAAGGAACTTAAAGGAACTTAAAGGAA 2223
DB |||||
QY 901 ATTCAGTGAAGAGTTGTTTTCAGATTGATCAAGCGGAGCCAAAGAACTATCAGTGGT 960
DB |||||
QY 2224 ATTCAGTGAAGAGTTGTTTTCAGATTGATCAAGCGGAGCCAAAGAACTATCAGTGGT 2283
DB |||||
QY 961 CAGAGTCCGGCGGAACTTCTCCGACCCAGGAGCCACACAACTGAGACACACAA 1020
DB |||||
QY 2284 CAGAGTCCGGCGGAACTTCTCCGACCCAGGAGCCACACAACTGAGACACACAA 2343
DB |||||
QY 1021 ATCATGGCTTCAGAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAAGGAGCT 1080
DB |||||
QY 2344 ATCATGGCTTCAGAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAAGGAGCT 2403
DB |||||
QY 1081 GCAGTGTGCAATCTAAACACCTTGCACAAATCTCCAGAGTCCCAATCCCTCAAC 1140
DB |||||
QY 2404 GCAGTGTGCAATCTAAACACCTTGCACAAATCTCCAGAGTCCCAATCCCTCAAC 2463
DB |||||
QY 1141 AAACAGGTCCGGACAAAGACCCCAATATACACCGGTGTATTAACCTTGACATCTCTGAG 1200
DB |||||
QY 2464 AAACAGGTCCGGACAAAGACCCCAATATACACCGGTGTATTAACCTTGACATCTCTGAG 2523
DB |||||
QY 1201 GCAACTCAAGTTGAAACAATCACCGCAGACAGACAGACAGACAGACAGACAGACAG 1260
DB |||||
QY 2524 GCAACTCAAGTTGAAACAATCACCGCAGACAGACAGACAGACAGACAGACAGACAG 2583
DB |||||
QY 1261 CCCTCTGCCAGACCGCAGCGGACCCCAAGAGAGAGAGTCCAGACAGACAGACAGACAG 1320
DB |||||
QY 2584 CCCTCTGCCAGACCGCAGCGGACCCCAAGAGAGAGAGTCCAGACAGACAGACAGACAG 2643
DB |||||
QY 1321 ACTGACTTCTGGACCCCGCAGCAGACCAAGTCCCAAGAGAGAGTCCAGACAGACAGACAG 1380
DB |||||
QY 2644 ACTGACTTCTGGACCCCGCAGCAGACCAAGTCCCAAGAGAGAGTCCAGACAGACAGACAG 2703
DB |||||
QY 1381 AACACACACACTCATCACCAGATACCGGAGAGAGTCCAGACAGACAGACAGACAGACAG 1440
DB |||||
QY 2704 AACACACACACTCATCACCAGATACCGGAGAGAGTCCAGACAGACAGACAGACAGACAG 2763
DB |||||
QY 1441 TTAATTACCAATACTATTGCTGAGTCCGAGGATGATCAGAGCGGGAGAGAACTCGA 1500
DB |||||
QY 2764 TTAATTACCAATACTATTGCTGAGTCCGAGGATGATCAGAGCGGGAGAGAACTCGA 2823
DB |||||
QY 1501 CGA 1503
DB |||||
QY 2824 AGA 2826
```

```
RESULT 14
US-09-337-946A-1
; Sequence 1, Application US/09337946A
; Patent No. US20020164582A1
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hart, Mary Katherine
; APPLICANT: Wilson, Julie A.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Ebola Virus Proteins Expressed from Venezuelan Equine Encephali
; TITLE OF INVENTION: Virus Replicons
; FILE REFERENCE: Army 144
; CURRENT APPLICATION NUMBER: US/09/337,946A
```

```
; CURRENT FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: US 60/091,403
; FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: IBM compatible, Word 97, Windows 95
; SEQ ID NO 1
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Ebola Zaire
; US-09-337-946A-1
```

```
Query Match 66.6%; Score 1499.8; DB 9; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 ATGGCGGTTACAGGAATATTGCAGTTTACCTCGTGATTCGATTCAAGAGGACATCATTTCTTT 60
DB |||||
QY 59 ATGGCGGTTACAGGAATATTGCAGTTTACCTCGTGATTCGATTCAAGAGGACATCATTTCTTT 118
DB |||||
QY 61 CTTTGGGTAATTTATCTTTTCCAAAGAAACATTTTCCATCCCATCTGGAGTCAATCAAT 120
DB |||||
QY 119 CTTTGGGTAATTTATCTTTTCCAAAGAAACATTTTCCATCCCATCTGGAGTCAATCAAT 178
DB |||||
QY 121 AGCACATTACAGGTTAGTGTGTCGACAACTAGTTTGTCTGACAAACTGTCTATCCACA 180
DB |||||
QY 179 AGCACATTACAGGTTAGTGTGTCGACAACTAGTTTGTCTGACAAACTGTCTATCCACA 238
DB |||||
QY 181 AATCAATTTAGATCAGTTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 240
DB |||||
QY 239 AATCAATTTAGATCAGTTTGGACTGAATCTCGAAGGGAATGAGTGGCAACTGACGTGCCA 298
DB |||||
QY 241 TCTGCAACTTAAAGAGTGGGGCTTTCAGGTCCGGTGTCCCAACCAAGGTGGTCAATTTATGAA 300
DB |||||
QY 299 TCTGCAACTTAAAGAGTGGGGCTTTCAGGTCCGGTGTCCCAACCAAGGTGGTCAATTTATGAA 358
DB |||||
QY 301 GCTGTGTAATGGGCTGAAAACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 360
DB |||||
QY 359 GCTGTGTAATGGGCTGAAAACTGCTACAACTTTGAAATCAAAAAACCTGACGGGAGTGAG 418
DB |||||
QY 361 TGCTTACACAGCAGCCAGACGGGATTCGGGCTTCCCGGTGCGGTGATGTGCACAAA 420
DB |||||
QY 419 TGTCTACACAGCAGCCAGACGGGATTCGGGCTTCCCGGTGCGGTGATGTGCACAAA 478
DB |||||
QY 421 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATTAAGAGGGTCTTTCTTC 480
DB |||||
QY 479 GTATCAGGAACGGGACCGGTGTCGGGAGACTTTGCTTCCATTAAGAGGGTCTTTCTTC 538
DB |||||
QY 481 CTGTATGATCGACTTGTCTTCCAGTTTCTACCGAGGACGACTTTCGCTGAAGGTGTC 540
DB |||||
QY 539 CTGTATGATCGACTTGTCTTCCAGTTTCTACCGAGGACGACTTTCGCTGAAGGTGTC 598
DB |||||
QY 541 GTTGCAATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 600
DB |||||
QY 599 GTTGCAATTTCTGATACTGCCCAAGCTAAGAGGACTTCTTCAGCTCAACCCCTTGAGA 658
DB |||||
QY 601 GAGCCGGTCAATGCAACCGGAGGACCGTCTAGTGGCTACTATTCTACCAATTAGATAT 660
DB |||||
QY 659 GAGCCGGTCAATGCAACCGGAGGACCGTCTAGTGGCTACTATTCTACCAATTAGATAT 718
DB |||||
QY 661 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
DB |||||
QY 719 CAGGCTACCGGTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 778
DB |||||
QY 721 TACGTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 780
DB |||||
QY 779 TACGTCCAACTTGAATCAAGATTTCACACACAGTTTCTGCTCCAGCTGAATGAGACAATA 838
DB |||||
QY 781 TATACAAGTGGGAAAGGAGCAATACCAGGGAATACTAATTTGGAGGTCAACCCCGAA 840
DB |||||
QY 839 TATACAAGTGGGAAAGGAGCAATACCAGGGAATACTAATTTGGAGGTCAACCCCGAA 898
DB |||||
QY 841 ATTGATACAACTCGGGAGTGGGCTTCTCGGAAACTTAAAGGAACTTAAAGGAACTTAAAGGAA 900
DB |||||
```

Db 899 ATTGATACACAAATCGGGAGTGGCCCTTCTGGGAACTAAAAAACCCTCACTAGAAA 958
Qy 901 ATTCCAGTGAAGATTGTCTTTTCAAGTTGATCAAAACGAGCCAAAAAATCATAGTGT 960
Db 959 ATTCCAGTGAAGATTGTCTTTTCAAGTTGATCAAAACGAGCCAAAAAATCATAGTGT 1018
Qy 961 CAGAGTCGGGCGGAATCTTCCGACCCAGGACCAACACAACTGAAGACACAAA 1020
Db 1019 CAGAGTCGGGCGGAATCTTCCGACCCAGGACCAACACAACTGAAGACACAAA 1078
Qy 1021 ATCATGGCTTCAGAAAAATCTCTGCAATGGTTCAAGTGCAAGTCAAGGAGGGAAGCT 1080
Db 1079 ATCATGGCTTCAGAAAAATCTCTGCAATGGTTCAAGTGCAAGTCAAGGAGGGAAGCT 1138
Qy 1081 GCAGTGTGCTACTAAACCCCTTCCCAATCTCCAGAGTCCCAATCCCTCAACAAC 1140
Db 1139 GCAGTGTGCTACTAAACCCCTTCCCAATCTCCAGAGTCCCAATCCCTCAACAAC 1198
Qy 1141 AAACGAGTCCGGACACAGACACCAATATACACCGGTGTATAACTTGACATCTCTGAG 1200
Db 1199 AAACGAGTCCGGACACAGACACCAATATACACCGGTGTATAACTTGACATCTCTGAG 1258
Qy 1201 GCAACTCAAGTTGAACAAATCATACCGCAGAAACAGACAGACAGACAGCCCTCCGACACT 1260
Db 1259 GCAACTCAAGTTGAACAAATCATACCGCAGAAACAGACAGACAGACAGCCCTCCGACACT 1318
Qy 1261 CCCTCTGCAACAGCGAGCGGACCCCAAAAGCAGAGAACCAACACAGGAGCAAGAGC 1320
Db 1319 CCCTCTGCAACAGCGAGCGGACCCCAAAAGCAGAGAACCAACACAGGAGCAAGAGC 1378
Qy 1321 ACTGACTCTCTGGACCCCGCACACACAGTCCCAACCCAGGAGACCGCTGGC 1380
Db 1379 ACTGACTCTCTGGACCCCGCACACAGTCCCAACCCAGGAGACCGCTGGC 1438
Qy 1381 AACAAACACACTCATCAACAGATACCGAGAGAGAGTGCCAGACAGCGGAAGCTAGGC 1440
Db 1439 AACAAACACACTCATCAACAGATACCGAGAGAGAGTGCCAGACAGCGGAAGCTAGGC 1498
Qy 1441 TTAATTAACAATATATTGCTGGATCGCAGACTGATCAGAGCGGAGAGAACTCGA 1500
Db 1499 TTAATTAACAATATATTGCTGGATCGCAGACTGATCAGAGCGGAGAGAACTCGA 1558
Qy 1501 CGA 1503
Db 1559 AGA 1561

RESULT 15

US-10-384-976-1

; Sequence 1, Application US/10384976
; Publication No. US20030224015A1
; GENERAL INFORMATION:
; APPLICANT: HART, MARY KATHERINE
; APPLICANT: WILSON, JULIE A.
; APPLICANT: PUSHKO, PETER
; APPLICANT: SMITH, JONATHAN F.
; APPLICANT: SCHWALJOHN, ALAN L.
; TITLE OF INVENTION: EBOLA PEPTIDES AND IMMUNOGENIC COMPOSITIONS CONTAINING
; TITLE OF INVENTION: SAME
; FILE REFERENCE: ARMY 144A
; CURRENT APPLICATION NUMBER: US/10/384,976
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 09/337,946
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2298
; TYPE: DNA
; ORGANISM: Ebola zaire
US-10-384-976-1

Query Match 66.6%; Score 1499.8; DB 17; Length 2298;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1501; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGGGCGTTACAGGAATATTGTCAGTTACCTCGTCGATCGATTCAGAGGACATCATTTCTTT 60
Db 59 ATGGGCGTTACAGGAATATTGTCAGTTACCTCGTCGATCGATTCAGAGGACATCATTTCTTT 118
Qy 61 CTTTGGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCACAAT 120
Db 119 CTTTGGGTAAATATCTCTTTTCCAAAGAACATTTTCCATCCCATCTGGAGTCATCCACAAT 178
Qy 121 AGCATTACAGGTTAGTGTGTCGACAACTAGTTTGTGTCGACAAACTGTGTATCTCATCCACA 180
Db 179 AGCATTACAGGTTAGTGTGTCGACAACTAGTTTGTGTCGACAAACTGTGTATCTCATCCACA 238
Qy 181 AATCAATTTGAGATCAGTTGGACTGAATCTCGAAGGGAATGGAGTGGCAACTGAGCTGCCA 240
Db 239 AATCAATTTGAGATCAGTTGGACTGAATCTCGAAGGGAATGGAGTGGCAACTGAGCTGCCA 298
Qy 241 TCTGCAACTAAAGATGGGCTTCAGGTCGGGTGTCGCCACCAAAAGGTGGTCAATATATGAA 300
Db 299 TCTGCAACTAAAGATGGGCTTCAGGTCGGGTGTCGCCACCAAAAGGTGGTCAATATATGAA 358
Qy 301 GCTGTTGAATGGGCTGAAAATCTGCTACAACTTGAATCAAAAACCTGACGGGAGTGAG 360
Db 359 GCTGTTGAATGGGCTGAAAATCTGCTACAACTTGAATCAAAAACCTGACGGGAGTGAG 418
Qy 361 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTCGCGGTATGTGCACAAA 420
Db 419 TGTCTACAGCAGCGCCAGACGGGATTCGGGGCTTCCCGGTCGCGGTATGTGCACAAA 478
Qy 421 GTATCAGGAAACGGACCGTGTGCGGAGACTTTCCTTCCATAAAGAGGGTGTCTTCTTC 480
Db 479 GTATCAGGAAACGGACCGTGTGCGGAGACTTTCCTTCCATAAAGAGGGTGTCTTCTTC 538
Qy 481 CTGTATGATCGACTTGTCTTCCAGAGTTATCTACCGAGGAAACGACTTTCGCTGAGAGGTGC 540
Db 539 CTGTATGATCGACTTGTCTTCCAGAGTTATCTACCGAGGAAACGACTTTCGCTGAGAGGTGC 598
Qy 541 GTTCATTTCTGATACTGCCCAAGCTTAAGAAGACTTCTTCAGTCAACCCCTTGAGA 600
Db 599 GTTCATTTCTGATACTGCCCAAGCTTAAGAAGACTTCTTCAGTCAACCCCTTGAGA 658
Qy 601 GAGCGGTCAATGCAACGGAGACCCGCTCTAGTGGCTACTATTCTACCAATTAGATAT 660
Db 659 GAGCGGTCAATGCAACGGAGACCCGCTCTAGTGGCTACTATTCTACCAATTAGATAT 718
Qy 661 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 720
Db 719 CAGGCTACCGGTTTTTGGAAACCAATGAGACAGAGTACTTGTTCGAGGTTGACAAATTTGACC 778
Qy 721 TACGTCCAACTTGAATCAAGATTCCACACAGTTTCTGCTCCAGCTGAATGAGACAAATA 780
Db 779 TACGTCCAACTTGAATCAAGATTCCACACAGTTTCTGCTCCAGCTGAATGAGACAAATA 838
Qy 781 TATCAAGTGGGAAAAGAGCAATACCAACGGGAAAATAATTTTGGAGGTCACACCCCGAA 840
Db 839 TATCAAGTGGGAAAAGAGCAATACCAACGGGAAAATAATTTTGGAGGTCACACCCCGAA 898
Qy 841 ATTGATACAAACATCGGGAGTGGGCTTCTGGGAAACTAAAAAACCCTCCTAGAAAAA 900
Db 899 ATTGATACAAACATCGGGAGTGGGCTTCTGGGAAACTAAAAAACCCTCCTAGAAAAA 958
Qy 901 ATTGCGAGTGAAGATTGTCTTTCACAGTTCTATCBAACGGGACCAAAACATCAGTGT 960
Db 959 ATTGCGAGTGAAGATTGTCTTTCACAGTTCTATCAAAACGGAGCCAAAAAATCATAGTGT 1018
Qy 961 CAGAGTCCGGCGCAACTTCTTCGACCCAGGAGCCAAACAACTGTAAGAGCCACAAA 1020
Db 1019 CAGAGTCCGGCGCAACTTCTTCGACCCAGGAGCCAAACAACTGTAAGAGCCACAAA 1078

Qy	1021	ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT	1080
Db	1079	ATCATGGCTTCAGAAAATTCCTCTGCAATGGTTCAAGTGCACAGTCAAGGAAGGGAAGCT	1138
Qy	1081	GCAGTGTGCGCATCTAACAAACCTTGGCCAAATCTCCACGAGTCCCAATCCCTCACAACC	1140
Db	1139	GCAGTGTGCGCATCTAACAAACCTTGGCCAAATCTCCACGAGTCCCAATCCCTCACAACC	1198
Qy	1141	AAACGAGGTCCGGACAAACAGCACCCATATACACCGTGTATATAAATTGACATCTCTGAG	1200
Db	1199	AAACGAGGTCCGGACAAACAGCACCCATATACACCGTGTATATAAATTGACATCTCTGAG	1258
Qy	1201	GCAACTCAAAGTTGAACAAATCACCGCAGAACAGACAAACAGACAGACAGCCTCCGACACT	1260
Db	1259	GCAACTCAAAGTTGAACAAATCACCGCAGAACAGACAAACAGACAGACAGCCTCCGACACT	1318
Qy	1261	CCCTCTGCCACGACCGCAGCCGGAACCCCAAAAGCAGAGAACCAACACGAGGGAAGAGC	1320
Db	1319	CCCTCTGCCACGACCGCAGCCGGAACCCCAAAAGCAGAGAACCAACACGAGGGAAGAGC	1378
Qy	1321	ACTGACTTCTTGACCCCGCCACCAACAAAGTCCCAAAACACAGCGGAGACCGCTGGC	1380
Db	1379	ACTGACTTCTTGACCCCGCCACCAACAAAGTCCCAAAACACAGCGGAGACCGCTGGC	1438
Qy	1381	AACAACAACTCATCAACCAAGATACCGGAGAGAGAGTGCCAGCAGCGGGAAGCTAGGC	1440
Db	1439	AACAACAACTCATCAACCAAGATACCGGAGAGAGAGTGCCAGCAGCGGGAAGCTAGGC	1498
Qy	1441	TTAATTACCAATACTATTGCTGGAGTCGAGGACTGATCACAGGGGGGAGAGAACTCGA	1500
Db	1499	TTAATTACCAATACTATTGCTGGAGTCGAGGACTGATCACAGGGGGGAGAGAACTCGA	1558
Qy	1501	CGA.1503	
Db	1559	AGA.1561	

Search completed: August 1, 2005, 20:40:55
Job time : 1421.31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:57:58 ; Search time 7912.14 Seconds

(without alignment)
11274.581 Million cell updates/sec

Title: US-10-066-506A-3

Perfect score: 1841
Sequence: 1 atgaagaccacatgttctc.....ttgtcttttagtgaattc 1841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sv.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1841	100.0	1841	6	AX717715	AX717715 Sequence
2	1308	71.0	2046	6	AX717717	AX717717 Sequence
3	1303.4	70.8	11460	6	AR279677	AX279677 Sequence
4	1303.4	70.8	19104	14	MVRBPCYC	Z12132 Marburg vir
5	1301.8	70.7	19112	14	AY430366	AY430366 Lake Vict
6	1301.8	70.7	19113	14	AY430365	AY430365 Lake Vict
7	1274	69.2	2051	6	AX717721	AX717721 Sequence
8	1151.4	62.5	2046	14	AF005735	AF005735 Marburg v
9	1148.2	62.4	2948	14	MAVSPAA	X68493 Marburg Vir
10	1148.2	62.4	19112	14	MVIRPR	Z29337 Marburg vir
11	1098.6	59.7	2046	14	AF005733	AF005733 Marburg v
12	1098.6	59.7	19151	14	AY358025	AY358025 Marburg v
13	780	42.4	2046	6	AX717719	AX717719 Sequence
14	780	42.4	2046	6	AX717725	AX717725 Sequence
15	777	42.2	2046	14	AF005734	AF005734 Marburg v
16	540	29.3	2039	6	AX717723	AX717723 Sequence
17	527.6	28.7	2406	14	EVU31033	U31033 Zaire Ebola
18	527.6	28.7	2408	14	EVU3187	U3187 Zaire Ebola
19	527.6	28.7	11742	14	EBORNA	L11365 Zaire Ebola

20	527.6	28.7	18959	14	AF086833	AF086833 Zaire Ebo
21	527.6	28.7	18959	14	AF272001	AF272001 Zaire Ebo
22	526	28.6	2298	6	AR404895	AR404895 Sequence
23	526	28.6	2298	6	AX092102	AX092102 Sequence
24	526	28.6	18959	14	AY142960	AY142960 Zaire Ebo
25	526	28.6	18960	14	AF499101	AF499101 Zaire Ebo
26	524.4	28.5	2408	14	EVU81161	EVU81161 Zaire Ebola
27	520.8	28.3	2408	14	EVU28077	EVU28077 Zaire Ebola
28	520.8	28.3	18961	14	AY354458	AY354458 Zaire ebo
29	516.4	28.0	2446	14	EVU77384	EVU77384 Zaire Ebola
30	514.8	28.0	2173	14	AY058898	AY058898 Zaire Ebo
31	511.6	27.8	2173	14	AY526098	AY526098 Zaire ebo
32	511.6	27.8	2173	14	AY526099	AY526099 Zaire ebo
33	511.6	27.8	2173	14	AY526100	AY526100 Zaire ebo
34	511.6	27.8	2173	14	AY526101	AY526101 Zaire ebo
35	511.6	27.8	2173	14	AY526102	AY526102 Zaire ebo
36	511.6	27.8	2173	14	AY526104	AY526104 Zaire ebo
37	510	27.7	1041	14	AY526103	AY526103 Zaire ebo
38	510	27.7	2173	14	AY526105	AY526105 Zaire ebo
39	324.2	17.6	2408	14	EVU28006	EVU28006 Cote d'Ivoi
40	286.6	15.6	2033	14	EVU23416	EVU23416 Reston Ebol
41	286.6	15.6	2033	14	EVU23417	EVU23417 Reston Ebol
42	286.6	15.6	2233	14	AF034645	AF034645 Ebola vir
43	286.6	15.6	2360	14	EVU23152	EVU23152 Reston Ebol
44	286.6	15.6	18891	14	AF522874	AF522874 Reston Eb
45	283.4	15.4	18890	14	AB050936	AB050936 Reston Eb

ALIGNMENTS

RESULT 1	AX717715	AX717715	1841 bp	DNA	linear	PAT 15-APR-2003
LOCUS	Sequence 3 from Patent WO02079239.					
DEFINITION	AX717715					
ACCESSION	AX717715.1	GI:29890726				
VERSION						
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	other sequences; artificial sequences.					
REFERENCE	1					
AUTHORS	Grogan,C.C., Hevey,M.C. and Schmaljohn,A.L.					
TITLE	Chimeric filovirus glycoprotein					
JOURNAL	Patent: WO 02079239-A 3 10-OCT-2002;					
FEATURES	U.S. Army Medical Research Institute of Infectious Diseases (US)					
source	Location/Qualifiers					
	1..1841					
	/organism="synthetic construct"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
	/note="Chimeric molecule between Marburg virus strain					
	Musoke Glycoprotein 1 and Ebola virus Zaire Mayinga strain					
	Glycoprotein 2"					

ORIGIN	Query Match	100.0%;	Score 1841;	DB 6;	Length 1841;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 1841;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC	60		
Db	1	ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC	60		
Qy	61	ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGGATTCCGTTATGTCCTCCGAAT	120		
Db	61	ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGGATTCCGTTATGTCCTCCGAAT	120		
Qy	121	CTCCAGAGACAGAGACGTCCTCTGATGGGATTACACTGAGTGGGCAAAAAGTTGCT	180		
Db	121	CTCCAGAGACAGAGAGCTCCATCTGATGGGATTACACTGAGTGGGCAAAAAGTTGCT	180		
Qy	181	GATTCCCTTTGGAGGCATCCAGGATGGGCTTTCAGGACAGGTGTACTCTCCCAAGAT	240		

181 GATTCCTCTTGGAGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
241 GTTGAGTACACAGAGGGGAGAGCCAAAACATGCTACAATATAAGTGTACAGGATCCC 300
241 GTTGAGTACACAGAGGGGAGAGCCAAAACATGCTACAATATAAGTGTACAGGATCCC 300
301 TCTGAAAAATCCTTGCTGTAGATCCTCTCAACACATCCGTGACTATCCGAAATGCAAA 360
301 TCTGAAAAATCCTTGCTGTAGATCCTCTCAACACATCCGTGACTATCCGAAATGCAAA 360
361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTATGG 420
361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTATGG 420
421 GGAGCATTTTTTCTGTATGATCGCATGCTCCCAACAAATGTACCGAGGCAAAATCTTC 480
421 GGAGCATTTTTTCTGTATGATCGCATGCTCCCAACAAATGTACCGAGGCAAAATCTTC 480
481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCG 540
481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCG 540
541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTAATAATATTTGGACAAGT 600
541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTAATAATAATTTGGACAAGT 600
601 ACTACGGAAACCAACGAATGACACTGATGTTTTCGGGCGCTTCAAGAAATACAAATCT 660
601 ACTACGGAAACCAACGAATGACACTGATGTTTTCGGGCGCTTCAAGAAATACAAATCT 660
661 ACAAGAAACCAACATGCTCGTCCAAAATACCTCCACACATGCGCCACAGCCGTCG 720
661 ACAAGAAACCAACATGCTCGTCCAAAATACCTCCACACATGCGCCACAGCCGTCG 720
721 GAGATCAAACTCACAAGCACCCCACTGATGCCAACAATCAATACCAAGCCCAAGC 780
721 GAGATCAAACTCACAAGCACCCCACTGATGCCAACAATCAATACCAAGCCCAAGC 780
781 AGTGATGATGAGGACCTCCGACATCCGCTCAGGTCGGAGAGCGAAACCCACACACA 840
781 AGTGATGATGAGGACCTCCGACATCCGCTCAGGTCGGAGAGCGAAACCCACACACA 840
841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTATCAACAAATGCGCCACTCCCTCACCA 900
841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTATCAACAAATGCGCCACTCCCTCACCA 900
901 CAACCAAGCACGCGCACAGCAAGGAGGAAACAAACAAACATTTCCCAAGATGCTGTGACT 960
901 CAACCAAGCACGCGCACAGCAAGGAGGAAACAAACAAACATTTCCCAAGATGCTGTGACT 960
961 GAACTAGACAAAATAACACAACTGACAAACCGTCCATGCCCTCATAACTACCACTACCACA 1020
961 GAACTAGACAAAATAACACAACTGACAAACCGTCCATGCCCTCATAACTACCACTACCACA 1020
1021 ATCTCTACTAAACAAACCTCCAAACAACTTTCAGCACTCTCTGCGACCATTAACAAC 1080
1021 ATCTCTACTAAACAAACCTCCAAACAACTTTCAGCACTCTCTGCGACCATTAACAAC 1080
1081 ACCACCAATGACAAACACAGACGACCAATCACTGAAAAATGAGCAAAACCAAGTCCCCCTCG 1140
1081 ACCACCAATGACAAACACAGACGACCAATCACTGAAAAATGAGCAAAACCAAGTCCCCCTCG 1140
1141 ATACCAACCTTGCTCCAGGGGAATCCCAACAGCAAGAGACGACGAGCAAAA 1200
1141 ATACCAACCTTGCTCCAGGGGAATCCCAACAGCAAGAGACGACGAGCAAAA 1200
1201 GGCCCCGACCAACAGCGCACCAACAGCAATGAGCATTTTACCAGTCTCTCCCCAC 1260
1201 GGCCCCGACCAACAGCGCACCAACAGCAATGAGCATTTTACCAGTCTCTCCCCAC 1260
1261 CCAGCTCGACTGCAACAACATCTGTATATTTTCAAGAAAGCGATTCGGCAATTTGTCAAT 1320

Db 1261 CCCAGCTCGATGCGCAACATCTTGTATATTTTCAAGAAAGCGATCGGCAATTTGTCAAT 1320
QY 1321 GCTCAACCAAAATCAACCCCTAAATTTTACTTACTGGACTACTCAGGATGAAGGTGCTGCA 1380
Db 1321 GCTCAACCAAAATCAACCCCTAAATTTTACTTACTGGACTACTCAGGATGAAGGTGCTGCA 1380
QY 1381 ATCGGACTGGCTGGATACCATATTTTCGGGCGAGCAGCGGAGGAATTTACATAGAGGG 1440
Db 1381 ATCGGACTGGCTGGATACCATATTTTCGGGCGAGCAGCGGAGGAATTTTACATAGAGGG 1440
QY 1441 CTAATGCACAATCAAGATGGTTTAAATCTGTGGTTGAGACAGCTGCGCAACGAGAGGACT 1500
Db 1441 CTAATGCACAATCAAGATGGTTTAAATCTGTGGTTGAGACAGCTGCGCAACGAGAGGACT 1500
QY 1501 CAAGCTCTTCAACTGTTCTCTGAGAGCCCAAACTGAGCTACGCACTTTTCAATCCTCAAC 1560
Db 1501 CAAGCTCTTCAACTGTTCTCTGAGAGCCCAAACTGAGCTACGCACTTTTCAATCCTCAAC 1560
QY 1561 CGTAAGGCAATGATTTCTTCTGAGCGATGGGCGGCAATGCGCAATTTCTGGAGCG 1620
Db 1561 CGTAAGGCAATGATTTCTTCTGAGCGATGGGCGGCAATGCGCAATTTCTGGAGCG 1620
QY 1621 GACTGCTATCGACCAATGATTTGGACCAAGACATTAACAGACAAAATTTGATCAGATT 1680
Db 1621 GACTGCTATCGACCAATGATTTGGACCAAGACATTAACAGACAAAATTTGATCAGATT 1680
QY 1681 ATTCATGATTTTGTGATAAAACCTTCCGACCAAGGGGCAATGACAAATTTGGTGACA 1740
Db 1681 ATTCATGATTTTGTGATAAAACCTTCCGACCAAGGGGCAATGACAAATTTGGTGACA 1740
QY 1741 GGATGAGCAATGATACCGGCAAGGATTTGAGTTTACAGGCGTTATTAATTCAGTTATC 1800
Db 1741 GGATGAGCAATGATACCGGCAAGGATTTGAGTTTACAGGCGTTATTAATTCAGTTATC 1800
QY 1801 GCTTTATTTCTGATATGCAAAATTTGCTTTTAGTTGAATTC 1841
Db 1801 GCTTTATTTCTGATATGCAAAATTTGCTTTTAGTTGAATTC 1841

RESULT 2
AX717717
LOCUS AX717717 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 5 from Patent WO02079239.
ACCESSION AX717717
VERSION AX717717.1 GI:29890727
KEYWORDS .
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Grogan,C.C.; Hevey,M.C. and Schmaljohn,A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 5 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source
1. .2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chimeric molecule between Marburg virus strain
Musoke Glycoprotein 1 and Marburg virus strain Raven
Glycoprotein 2"

ORIGIN
Query Match 71.0%; Score 1308; DB 6; Length 2046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAATCTCCCC 60
Db 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAATCTCCCC 60
QY 61 ATTTTAGATAGTAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGA 120

Db 61 ATTTAGAGATAGCTAGTAATATCAACCCCAATGTGGATTCGGTATGCTCCGGAAT 120
Qy 121 CTCAGAGACAGAAAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCAGAGACAGAAAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Qy 181 GATTCCCTTTGGAGGATCCAGGATGGGCTTTCCAGACAGGTGTACCTCCAGAAAT 240
Db 181 GATTCCCTTTGGAGGATCCAGGATGGGCTTTCCAGACAGGTGTACCTCCAGAAAT 240
Qy 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAAGGATCCC 300
Db 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAAGGATCCC 300
Qy 301 TCTGAAATCTCTGCTGTATGATCCTCTCAACCAATCCGTGACTATCCGAAATGCAAA 360
Db 301 TCTGAAATCTCTGCTGTATGATCCTCTCAACCAATCCGTGACTATCCGAAATGCAAA 360
Qy 361 ACTATCCATCATATTCAGAGGTCAAAACCTCTGATGACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTATCCATCATATTCAGAGGTCAAAACCTCTGATGACAGGGGATCGCCCTTCATTTATGG 420
Qy 421 GGAGCATTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAAAGTCTTC 480
Db 421 GGAGCATTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAAAGTCTTC 480
Qy 481 ACTGAAGGGAACATAGCAGCTATGATTTGCAATAGACAGTGCACAAAATGATTTTCTCG 540
Db 481 ACTGAAGGGAACATAGCAGCTATGATTTGCAATAGACAGTGCACAAAATGATTTTCTCG 540
Qy 541 CGGCAAGGACAGAGGTACCGTCAATATGAAATCTGACTTCTAATAAATATTTGGACAAGT 600
Db 541 CGGCAAGGACAGAGGTACCGTCAATATGAAATCTGACTTCTAATAAATATTTGGACAAGT 600
Qy 601 ACTAAGGGAACCAACGAAATGACACTGGATGTTTCGGGCTCTTCAAGAAATACAAATCT 660
Db 601 ACTAAGGGAACCAACGAAATGACACTGGATGTTTCGGGCTCTTCAAGAAATACAAATCT 660
Qy 661 ACAAGAGACCAACATGTGCTCCGTCCAAATATACCTCCACCACTCCGACAGCCCGTCCG 720
Db 661 ACAAGAGACCAACATGTGCTCCGTCCAAATATACCTCCACCACTCCGACAGCCCGTCCG 720
Qy 721 GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCAATACCGAGCCCAAGC 780
Db 721 GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCAATACCGAGCCCAAGC 780
Qy 781 AGTGATGAGGACCTCGCAACATCCGCTCAGGCTCGGAGACGAGAACCCCAACA 840
Db 781 AGTGATGAGGACCTCGCAACATCCGCTCAGGCTCGGAGACGAGAACCCCAACA 840
Qy 841 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCATCAACAAATGCCACCTCCCTCACC 900
Db 841 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCATCAACAAATGCCACCTCCCTCACC 900
Qy 901 CAACCAAGCAGCCACAGCAGAGGAAACAAACCAATTCCTCAAGATGCTGTGACT 960
Db 901 CAACCAAGCAGCCACAGCAGAGGAAACAAACCAATTCCTCAAGATGCTGTGACT 960
Qy 961 GAACTAGACAAAATTAACAACACTGCAACCGTCCATGCTCCCTCATTAACACTACCA 1020
Db 961 GAACTAGACAAAATTAACAACACTGCAACCGTCCATGCTCCCTCATTAACACTACCA 1020
Qy 1021 ATCTCTACTTAAACAACCTTCAACCAACTTTCAGCACTCTCTGCACTTACAAAC 1080
Db 1021 ATCTCTACTTAAACAACCTTCAACCAACTTTCAGCACTCTCTGCACTTACAAAC 1080
Qy 1081 ACCACCAATGACAACAACAGAGCAAAATCACTGAAATGAGCAAAACCAAGTCCCTCTCG 1140
Db 1081 ACCACCAATGACAACAACAGAGCAAAATCACTGAAATGAGCAAAACCAAGTCCCTCTCG 1140
Qy 1141 ATAAACACCTGCTTCAACGGGAAATCCCAACACAGGAAAGACACCGAGGCAAAA 1200

Db 1141 ATAACAACCTGCTCCACCGGAAATCCCAACAGACAGACAGACAGCAAAAA 1200
Qy 1201 GGCCCCCGCCACACCGCACAAAACAGCAAAATGAGATTTCCACAGTCTCTCCCCACC 1260
Db 1201 GGCCCCCGCCACACCGCACAAAACAGCAAAATGAGATTTCCACAGTCTCTCCCCACC 1260
Qy 1261 CCCAGCTCGACTGCACACATCTTGTATATTTTCAGAGAAAGCGATCG 1308
Db 1261 CCCAGCTCGACTGCACACATCTTGTATATTTTCAGAGAAAGCGATCG 1308
RESULT 3
AR279677 11460 bp DNA linear PAT 10-APR-2003
LOCUS AR279677 Sequence 1 from patent US 6517842.
DEFINITION AR279677
ACCESSION AR279677
VERSION AR279677.1 GI:29714600
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11460)
AUTHORS Hevey, M.C., Negley, D.L., Pushko, P., Smith, J.F. and Schmaljohn, A.L.
TITLE Marburg virus vaccines
JOURNAL Patent: US 6517842-A 1 11-FEB-2003;
FEATURES Location/Qualifiers
source 1..11460
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 70.8%; Score 1303.4; DB 6; Length 11460;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAAATCTCCC 60
Db 5940 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAAATCTCCC 5999
Qy 61 ATTTTAGAGATAGCTAGTAATATAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 6000 ATTTTAGAGATAGCTAGTAATATAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 6059
Qy 121 CTCAGAGACAGAAAGAGCTCCATCTGATGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 6060 CTCAGAGACAGAAAGAGCTCCATCTGATGGATTCACACTGAGTGGGCAAAAAGTTGCT 6119
Qy 181 GATTCCCTTTGGAGGATCCAGGATGGGCTTTCCAGACAGGTGTACCTCCAGAAAT 240
Db 6120 GATTCCCTTTGGAGGATCCAGGATGGGCTTTCCAGACAGGTGTACCTCCAGAAAT 6179
Qy 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAAGGATCCC 300
Db 6180 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAAGGATCCC 6239
Qy 301 TCTGAAATCTCTGCTGTATGATCCTCTCAACCAATCCGTGACTATCCGAAATGCAAA 360
Db 6240 TCTGAAATCTCTGCTGTATGATCCTCTCAACCAATCCGTGACTATCCGAAATGCAAA 6299
Qy 361 ACTATCCATCATATTCAGAGGTCAAAACCTCTGATGACAGGGGATCGCCCTTCATTTATGG 420
Db 6300 ACTATCCATCATATTCAGAGGTCAAAACCTCTGATGACAGGGGATCGCCCTTCATTTATGG 6359
Qy 421 GGAGCATTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAAAGTCTTC 480
Db 6360 GGAGCATTTTCTGTATGATCGCATTCCTCCCAACAATGTACCGAGGCAAAAGTCTTC 6419
Qy 481 ACTGAAGGGAACATAGCAGCTATGATTTGCAATAGACAGTGCACAAAATGATTTTCTCG 540
Db 6420 ACTGAAGGGAACATAGCAGCTATGATTTGCAATAGACAGTGCACAAAATGATTTTCTCG 6479
Qy 541 CGGCAAGGACAGAGGTACCGTCAATATGAAATCTGACTTCTAATAAATATTTGGACAAGT 600

Db	6480	CGGCAAGGACAAGGATGACCGTCAATGAAATCTGACTTCTACTAATAAATATTGGCAAGT	6539	AUTHORS	Muhlberger, E., Sanchez, A., Randolph, A., Will, C., Kiley, M.P., Klenk, H.D. and Feldmann, H.
Qy	601	AGTAACGGAAACGAAACGAATGACACTGATGTTTCGGCGCTCTTCAAGAAATACAAATCT	660	TITLE	The nucleotide sequence of the L gene of Marburg virus, a filovirus: homologies with paramyxoviruses and rhabdoviruses
Db	6540	AGTAACGGAAACGAAACGAATGACACTGATGTTTCGGCGCTCTTCAAGAAATACAAATCT	6599	JOURNAL	Virology 187 (2), 534-547 (1992)
Qy	661	ACAAAGAACCAAAACATGTCTCGTCCAAATAATCTTCCACCACTGCGCCACAGCCCGTCCG	720	MEDLINE	92188528
Db	6600	ACAAAGAACCAAAACATGTCTCGTCCAAATAATCTTCCACCACTGCGCCACAGCCCGTCCG	6659	PUBMED	1546452
Qy	721	GAGATCAAACTCACAAGCAACCCCACTGATGCCCAAACTCAATACCAAGCCCAAGC	780	AUTHORS	Feldmann, H., Muhlberger, E., Randolph, A., Will, C., Kiley, M.P., Sanchez, A. and Klenk, H.D.
Db	6660	GAGATCAAACTCACAAGCAACCCCACTGATGCCCAAACTCAATACCAAGCCCAAGC	6719	TITLE	Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle
Qy	781	AGTGATGATGAGGACCTCCCAACATCCGGCTCAGGGTCGGAGAGACGAGAACCCACACACA	840	JOURNAL	Virus Res. 24 (1), 1-19 (1992)
Db	6720	AGTGATGATGAGGACCTCCCAACATCCGGCTCAGGGTCGGAGAGACGAGAACCCACACACA	840	MEDLINE	92327834
Qy	841	ATTCTGTATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA	900	PUBMED	1626422
Db	6780	ATTCTGTATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA	6839	REMARK	(sites)
Qy	901	CAACCAAGCAGCCGACGAGGAGGAGAAACACACAAACCATTCACCAAGATCTGTGACT	960	AUTHORS	Will, C., Muhlberger, E., Linder, D., Slenczka, W., Klenk, H.D. and Feldmann, H.
Db	6840	CAACCAAGCAGCCGACGAGGAGGAGAAACACACAAACCATTCACCAAGATCTGTGACT	6899	TITLE	Marburg virus gene 4 encodes the virion membrane protein, a type I transmembrane glycoprotein
Qy	961	GAACTAGACAAAATAACAACTGACAAACCGTCCATGCCCGCTCATAACTACTACCA	1020	JOURNAL	J. Virol. 67 (3), 1203-1210 (1993)
Db	6900	GAACTAGACAAAATAACAACTGACAAACCGTCCATGCCCGCTCATAACTACTACCA	6959	MEDLINE	93172334
Qy	1021	ATCTCTACTAAACAACCTCCAAACAACTTTCAGCACTCTCTCTGCACTTACCAAAAC	1080	PUBMED	8437211
Db	6960	ATCTCTACTAAACAACCTCCAAACAACTTTCAGCACTCTCTCTGCACTTACCAAAAC	7019	REFERENCE	5
Qy	1081	ACCACCAATGACACACAGAGCAACATCACTGAAATGACAAACGATGCCCGCTCG	1140	AUTHORS	Feldmann, H.
Db	7020	ACCACCAATGACACACAGAGCAACATCACTGAAATGACAAACGATGCCCGCTCG	7079	TITLE	Direct Submission
Qy	1141	ATAACAACCTGCTCCCAACGGGAATCCACACAGAAAGACGACGAGCAAGAAA	1200	JOURNAL	Submitted (20-SEP-1994) Feldmann H., Zentrum fuer Hygiene der Philipps-Universitaet, Institut fuer Virologie, Robert-Koch-Str. 17, Marburg, Germany, 3550
Db	7080	ATAACAACCTGCTCCCAACGGGAATCCACACAGAAAGACGACGAGCAAGAAA	7139	REMARK	revised by [5] MAT
Qy	1201	GGCCCCGACCAACGGCACCACCAACGAAATGAGCAATTTACCAAGTCTCTCCCCCACC	1260	AUTHORS	6 (bases 1 to 19104)
Db	7140	GGCCCCGACCAACGGCACCACCAACGAAATGAGCAATTTACCAAGTCTCTCCCCCACC	7199	TITLE	Feldmann, H.
Qy	1261	CCAGCTCAGCTGCACAACTCTGTATATTTTCAGAGAAAGCGA	1305	JOURNAL	Submitted (20-SEP-1994) Feldmann H., Zentrum fuer Hygiene der Philipps-Universitaet, Institut fuer Virologie, Robert-Koch-Str. 17, Marburg, Germany, 3550
Db	7200	CCAGCTCAGCTGCACAACTCTGTATATTTTCAGAGAAAGCGA	7244	COMMENT	On Sep 22, 1994 this sequence version replaced gi:60623.
RESULT 4				FEATURES	Location/Qualifiers
MVREPCYC				source	1..19104
LOCUS				gene	/organism="Lake Victoria marburgvirus"
DEFINITION				mRNA	/mol_type="genomic DNA"
					/strain="Musoke"
					/db_xref="taxon:11269"
					47..2843
					/gene="np"
					47..2843
					/gene="np"
					/standard_name="gene one"
					/citation=[1]
					/citation=[3]
					/evidence=experimental
					104..2182
					/gene="np"
					/standard_name="np"
					/citation=[1]
					/citation=[3]
					/codon_start=1
					/evidence=experimental
					/product="nucleoprotein"
					/protein_id="CAA78114.1"
					/db_xref="GI:60624"
					/db_xref="GOA:P27588"
					/db_xref="UniProt/Swiss-Prot:P27588"
					/translation="MDLSHLSLELGTPTAPHRNKKVILFDTHQVHSICNQLIDAINS
					GIDGLLEGGLLTLCVHHYNSDKDNTSPVAKLRDAGTEFVDVINKNADATRELDV
					SPNEPHYPLIALKLTSTESQRIQGLFSLFKPLVGVGRASIEKALRQTV
					HQEOQIVTPNHLTTHGMKVFILGRSFLIKFVLHQVNLVTHDAYDSIISNV
					QTRFSGLLIVKLEFLQKDSQVTLHPLVRSKVKNEVASFKQALSNLARHGEYA
					PFARVNLGINNLEHGLYPOLSAALGATAHGSTLAGVNVGEQVQQLREAAHDAEV
					KLQRHHEQIEQAIADDEERKILEQFLQKTEITHSOTLAVLSQKREKLALAEIE
					NNIVEDQGFQSQNRVDSFLNDPTTPEVTQARPMNRPTALPPFPVDDKIERHSTEDS
					SSSSSFVLDNDFALLNDEDTLDDSVMIPTGTSREFQGIPEPPRQSQDLNNSQKQE
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
PUBMED					
REFERENCE					

DESTNRIKKOFLRYOELPPVOEDDESEYTTDSQESILTDQSGSDNEQGVLDLPPPLVYAOE
KRQDIIQHPAANQDPFGSIGVNGVDIIIEPINSPPSPAPQEDTRRAYELSDQFTN
DDNQONPQORVVITKGRTEFLYPNDLLQTNPPESLITALVBEYQNPVSAKSLDQWPD
MSFDDGDMLR
/gene="vp35"
/db_xref="GI:4409
/standard_name="gene two"
/citation=[3]
/evidence=experimental
2944..3933
/gene="vp35"
/standard_name="vp35"
/codon_start=1
/evidence=experimental
/product="vp35"
/protein_id="CAA78115.1"
/db_xref="GI:60625"
/db_xref="UniProt/Swiss-Prot:P35259"
/translation="NMDSYMQQVSEGLMGKVPIDQVFGANPLEKLYKRRKPKGTG
LQCSPLMSKATSDTDIIIVKRTILADLLIPINROI SDIQTSLSEVTVRVHEIER
LOHEITPVLMKGRTEALISKGMSMLAKYDHLVISTGRTTAPAAAFDAYLNEHGVPP
QPAIKDGLVQAQCSKGTWKQNTTDAADKMSKVLSESETFKPNLSAKDLALLF
TLPNGNTPFHILAQVLSKYGSGAFDAPHQIISGENAQAALTRLSTPDAF
LGVPPVIRVKNFQTVPRPSQKSLRAVPNPETIDKGWVCVYSSSEQETRAUKI"
4412..5818
/gene="vp40"
4412..5818
/gene="vp40"
/standard_name="gene three"
/citation=[3]
/evidence=experimental
4567..5478
/gene="vp40"
/standard_name="vp40"
/codon_start=1
/evidence=experimental
/product="vp40"
/protein_id="CAA78116.1"
/db_xref="GI:1552357"
/db_xref="GOA:P35260"
/translation="MASSSNVTYMQYLNPPYADHGANQLIPADQLSNOQGITPNVY
GDLNDDQFGKGVCHAPTLEAIDISAVNERTVKGVPALPLGIMSNFEPYLAHTVAA
LITGSYTTQFTHNGQKFRVNRLTGIPAHPLMLRGNQAFIQNNWVIPRNFSTNOR
TYNLNLVLSVQKLPDDARWRSKDKLIGNTHMPAVSIHPNLPPIVLPVTKQAYRQHK
NPNNGPLLAISGILHQLRVERKVPKTSLSLFRISLPADMFVSVEGMMKKRGENSPVYFQ
APENPFLNGFNNRQVFLAYNPVLSAV"
5822..8666
/gene="gp"
5822..8666
/gene="gp"
/standard_name="gene four"
/citation=[3]
/evidence=experimental
5940..7985
/gene="gp"
/standard_name="gp"
/codon_start=1
/evidence=experimental
/product="glycoprotein"
/protein_id="CAA78117.1"
/db_xref="GI:60627"
/db_xref="GOA:P35253"
/translation="NKTCTFLISLILIOGKNLPILIEASNNQPNQVNSVCGSLTQKT
EDVHLMGFTLSGQKVDSPLEASKEWAFRTGVPPKNVEYTEGEAKTCYNI SVTPDSG
KSLILDPPNIRDPKCTTIHQHQQONPHQAQIALHGWAFPLDRIASTNWRKVP
TEGNTAAMLVNKTVHKMIFSRQGGYRHNMLTSTNKYWTSSNGTNDTIDTCGCGALQEY

gene
mRNA
CDS

gene
mRNA
CDS

gene
mRNA
CDS

NSTKNQTCAPSKIPPLPTARPEIKLSTPTDATKLTNTDPSDDDELATSGSGSGER
EPHTTDAVTKRGLSSMPTPTSPQSGNNTHNSQAVTELDKNNTIAQSPMP
PHNTTISTNNTSKHNFSTLSAPLQNTNDNTOSTITENEQTSAPSIITLPTGNPTT
AKSTSSKGAAPTAPNTTNEHFTSPPTPSAQHLVFRKRRIILWREGMFPFLDG
LINAPIDPVPNTKTIIDESSSSGASAEEDQHASPNISLTLSPFNINENTAYSGEN
ENDCAELRTWSVOEDDLAAGLSWIPFPGPIEGLYAVLTKNNNLVCLRLRLANOT
AKSLELLRVTTERTFSLINRHAIDFLLPFGSTCKVLGPDCCIGIEDLSKNISEQI
DQIKDQKQEGTGMGLGKWWTSOWGLTNLGLILLLSIAVLIALSCICRIKTYIG"
8762..10009
/gene="vp30"
8762..10009
/gene="vp30"
/standard_name="gene five"
/citation=[3]
/evidence=experimental
8864..9697
/gene="vp30"
/standard_name="vp30"
/codon_start=1
/evidence=experimental
/product="vp30"
/protein_id="CAA78118.1"
/db_xref="GI:541782"
/db_xref="GOA:P35258"
/translation="MQPGRGRSTRNHQVTIYHETOLPKSPHYTHPHRARNMSST
RSSAESPTNHIIPRPPSTFNLSPKPPPPKDMCRNMKIGLPCADPTCNRDHLDNLT
NRELLLMARKMLPNTDKTPMODCGSPSLSKGLSKDQKQVLTDLNGLHLSY
LHRSEIGNWMBHRLAALSITCAGIRKTNRSILINTMTLHMHENLPODQNGVIKOTYT
GIHLDKGQGFEMALWQWDRKRSISLFFVQAALYVMNIPCESSISVQASYESFYSSKS
R"
9990..11278
/gene="vp24"
Query Match 70.8%; Score 1303.4; DB 14; Length 19104;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAAGACCATGTTTCCTTATCAGTCTTATCTTAATCAAGGGACAAAAATCTCCC 60
DB 5940 ATGAAGACCATGTTTCCTTATCAGTCTTATCTTAATCAAGGGACAAAAATCTCCC 5999
QY 61 ATTTAGAGTAGCTAGTAATAACACCCAAAATGGATTCGGTATGCTCCGGAAT 120
DB 6000 ATTTAGAGTAGCTAGTAATAACACCCAAAATGGATTCGGTATGCTCCGGAAT 6059
QY 121 CTCAGAGACAGAGAGCGTCCATCTGATGGATTTCACATGAGTGGGCAAAAAGTTGCT 180
DB 6060 CTCAGAGACAGAGAGCGTCCATCTGATGGATTTCACATGAGTGGGCAAAAAGTTGCT 6119
QY 181 GATTCCTCTTGGAGGCATCAAGCGATGGGCTTTTCAGACAGAGTGTACCTCCCAAGAA 240
DB 6120 GATTCCTCTTGGAGGCATCAAGCGATGGGCTTTTCAGACAGAGTGTACCTCCCAAGAA 6179
QY 241 GTTGAAGTACAGAGGGGGAGAAAGCCAAAACATGCTACAATATAAGTGTACGATGCC 300
DB 6180 GTTGAAGTACAGAGGGGGAGAAAGCCAAAACATGCTACAATATAAGTGTACGATGCC 6239
QY 301 TCTGGAAATCCTTGTCTGTAGATCCTCTACCAACATCCGCTACTATCCGAAATGCAAA 360
DB 6240 TCTGGAAATCCTTGTCTGTAGATCCTCTACCAACATCCGCTACTATCCGAAATGCAAA 6299
QY 361 ACTATCCATCATATTCAGAGTCAAAAACCTCATGACAGGGGATCGCCCTTCATTATG 420
DB 6300 ACTATCCATCATATTCAGAGTCAAAAACCTCATGACAGGGGATCGCCCTTCATTATG 6359
QY 421 GGAGCATTTTTCTGTATGATGCGATTCGCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
DB 6360 GGAGCATTTTTCTGTATGATGCGATTCGCTCCACAAATGTACCGAGGCAAAAGTCTTC 6419
QY 481 ACTGAGGGAACATAGGAGCTATGATTTGTCATATAGACAGTGCACAAATGATTTCTCG 540

Db	6420	ACTGAAGGGAACATAGCAGCTATGATTGTCCAATAAGACAGTGCACAAAATGATTTTCTCG	6479		
QY	541	CGGCAAGGACAAAGGGTACCGTCAATATGAATCTGACTTCTACTAATAAATATTTGGACAAGT	600		
Db	6480	CGGCAAGGACAAAGGGTACCGTCAATATGAATCTGACTTCTACTAATAAATATTTGGACAAGT	6539		
QY	601	AGTAACGGAAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAATACAAATTC	660		
Db	6540	AGTAACGGAAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAATACAAATTC	6599		
QY	661	ACAAAGAACCAACATGTGCTCCGTCGACCAATACCTCCACCTGCCCCCAGCCCCGTC	720		
Db	6600	ACAAAGAACCAACATGTGCTCCGTCGACCAATACCTCCACCTGCCCCCAGCCCCGTC	6659		
QY	721	GAGATCAAACTCACAAGCACCCCAACTGATGTCACCAACCAACTCAATACCAACGACCCCAAGC	780		
Db	6660	GAGATCAAACTCACAAGCACCCCAACTGATGTCACCAACCAACTCAATACCAACGACCCCAAGC	6719		
QY	781	AGTGATGATGAGGACCTTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAAACCCCAACACA	840		
Db	6720	AGTGATGATGAGGACCTTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAAACCCCAACACA	6779		
QY	841	ACTTCTGATCGGTCACCAAGCAAGGGCTTTCATCAACATGCCACCACTCCCTCACCA	900		
Db	6780	ACTTCTGATCGGTCACCAAGCAAGGGCTTTCATCAACATGCCACCACTCCCTCACCA	6839		
QY	901	CAACCAAGCACGCCACAGCAAGGAGGAAACAAACAACCAATTCCTCCCAAGATGCTGTGACT	960		
Db	6840	CAACCAAGCACGCCACAGCAAGGAGGAAACAAACAACCAATTCCTCCCAAGATGCTGTGACT	6899		
QY	961	GAACTAGACAAAAATAACAACAATGACAAACCGTCCATGCCCCCTCATATACCTACCA	1020		
Db	6900	GAACTAGACAAAAATAACAACAATGACAAACCGTCCATGCCCCCTCATATACCTACCA	6959		
QY	1021	ATCTCTACTAACAACACCTCCAAACACAACCTTCAGCACTCTCTCTGCACCATTAACAAAC	1080		
Db	6960	ATCTCTACTAACAACACCTCCAAACACAACCTTCAGCACTCTCTCTGCACCATTAACAAAC	7019		
QY	1081	ACCACCAATGACAAACACAGAGACACAATCACTGAAATGACGAAACGATGCCCCCTCG	1140		
Db	7020	ACCACCAATGACAAACACAGAGACACAATCACTGAAATGACGAAACGATGCCCCCTCG	7079		
QY	1141	ATAACAACCTCGCTCCAAACGGGAAATCCCAACACAGCAAGAGAGCAACGACGCAAAAA	1200		
Db	7080	ATAACAACCTCGCTCCAAACGGGAAATCCCAACACAGCAAGAGAGCAACGACGCAAAAA	7139		
QY	1201	GGCCCCGCAACAGGCAACCAACAGCAAAATGAGCATTTTCCAGTCTCTCCCCCACC	1260		
Db	7140	GGCCCCGCAACAGGCAACCAACAGCAAAATGAGCATTTTCCAGTCTCTCCCCCACC	7199		
QY	1261	CCAGCTCGACTGCACAACTCTTGATATTTTCAGAGAAAGCGA	1305		
Db	7200	CCAGCTCGACTGCACAACTCTTGATATTTTCAGAGAAAGCGA	7244		
RESULT 5	AY430366	19112 bp	RNA	linear	VRL 06-FEB-2004
LOCUS	AY430366	Lake Victoria marburgvirus strain pp4 guinea pig nonlethal variant, complete genome.			
ACCESSION	AY430366				
VERSION	AY430366.1	GI:40388379			
KEYWORDS	Lake Victoria marburgvirus				
SOURCE	Lake Victoria marburgvirus				
ORGANISM	Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.				
REFERENCE	1	(bases 1 to 19112)			
AUTHORS	Hevey, M., Negley, D., Geisbert, J., Jahrling, P. and Schmaljohn, A.				
TITLE	Antigenicity and vaccine potential of Marburg virus glycoprotein expressed by baculovirus recombinants				
JOURNAL	Virology 239 (1), 206-216 (1997)				
MEDLINE	98087840				

PUBMED	9426460	2	(bases 1 to 19112)				
REFERENCE	Chain, P.S.G., Malfatti, S.A., Hajjaj, A., Vergez, L.M., Do, L.H.,						
AUTHORS	Smith, K.L. and McCready, P.M.						
TITLE	Direct Submission						
JOURNAL	Submitted (09-OCT-2003) Viral Sequencing Group, Lawrence Livermore						
REFERENCE	National laboratory, 7000 East Avenue, Livermore, CA 94550, USA						
AUTHORS	3 (bases 1 to 19112)						
TITLE	Ichou, M.A., Paragas, J., Jahrling, P.B., Ibrahim, M.S., Lofte, L.,						
JOURNAL	Hevey, M. and Schmaljohn, A.						
FEATURES	Direct Submission						
source	Submitted (09-OCT-2003) Viral Genomics Group, USAMRIID, Fort						
gene	Detrick, MD 21702, USA						
CDS	Location/Qualifiers						
	1. .19112						
	/organism="Lake Victoria marburgvirus"						
	/mol_type="genomic RNA"						
	/strain="pp4 guinea pig nonlethal variant"						
	/isolate="Musoke Kenya, 1980"						
	/db_xref="taxon:11269"						
	104..2191						
	/gene="NP"						
	104..2191						
	/gene="NP"						
	/note="nucleoprotein"						
	/codon_start=1						
	/product="NP protein"						
	/protein_id="AAR85453.1"						
	/db_xref="GI:40388380"						
	/translation="MDLHLLLELGTGTAPHRNKKVILFDTNHQVSCNQIIDAIS						
	GIDGDLLEGGLLTCVHYNSDKFNTSPIAKYLRDAGYEVDFIKNAATRELDV						
	IPNEPHYSPIILALKTLESTESQRIGLFLSFLPLKLVGDRASIEKALROVTY						
	HOEGIVTYPNHMLTGTGMKVIPIGLIRSSFLKFLVLIHQGNLVTGHDAYSIIISNV						
	GOTFSGLLIVKTVLETLQKTSGLTLPVLRVTSKVKNEVASFKQALNSLARHGEYA						
	PFAVNLISGNLNLHGLYFOLSAIALGVATAHGSTLAGVNVGEOYQQLREAHDAEV						
	KLQRHHEQEIQAIEDDEERKILEQFLQKTEITHSQTALVLSQREKLARAEIE						
	NNIVDEGQKFSQNRVQSFLNDPTVEVTVOARPMNRPTALPPVDDDKIEHSTEDS						
	SSSSVFDLNDPFLALNEDEDLDDSVIMPCTTSREFGQIPEPRQSDQLNSQKQE						
	DESTNPKFLQYLOELPPVOEDDESEYTTDSQESIDQSGDNEQGVLDLPPPLYAQE						
	KRODPIOHPAANPODPFGSIGDVNGDILEPIRSPSPSAPOEDTKMREAYELSPDTN						
	DEDQIQHPAANPODPFGSIGDVNGDILEPIRSPSPSAPOEDTKMREAYELSPDTN						
	MSFDERHRVAVNL"						
gene	2945..3934						
CDS	2945..3934						
	/gene="VP35"						
	/note="minor component of RNA-dependent RNA-polymerase"						
	/codon_start=1						
	/product="VP35 protein"						
	/protein_id="AAR85454.1"						
	/db_xref="GI:40388381"						
	/translation="MWDSSYMQVSEGLMTGKVPIDOVFGANPLKLYKRRKPKCTVG						
	LQCSPLMSKATSDDDIWDQLIVKRTLADLLIPINROIISDIQSTLSEVTTTRVHEIER						
	QLHEITPVLMKGRTELEISKENSEMLAKIDYDLIVISTGRITAPAAAFDAYLNEHGVPPP						
	QPAIFKDLGVAQAQCSKGTMYKNATDAAKMSKVLSESETFSKENPLSAKDLALLLF						
	THLPENNTPHILAIQVLSKIAKYSKGSQAFDAPHQILSEGENAQAALTRLSRTFDAF						
	LGVVPPVIRVKNFQTPRCPQSKSLRAVPNPNTIDKGWVCVYSSEQGETRALKI"						
gene	4568..5479						
CDS	4568..5479						
	/gene="VP40"						
	/note="matrix protein"						
	/codon_start=1						
	/product="VP40 protein"						
	/protein_id="AAR85455.1"						
	/db_xref="GI:40388382"						
	/translation="MASSSNNTYNTYQYLNPPFYADHGANQLIPADQLNSQQGQITPNVY						
	GDLLDDQFGKNVCHAFTELEAIIISAYNTRTVKGVPAWFLGINSFEYPLAHVAA						
	LLTGSYTTIQTTHNGQKVRNRLGTGIPAHPLRLREGNQAFIQNVIPIRNFSTNQK						
	TYNLTNVLVSQKLDDAWRPSKDKLIGNTWHPAVSIHPNLPIIPLPTVKQQAQYQHK						
	NPNNGPLLAISGILHQLRVEKVPKTSLSFRISLPADMFVSKGMMKKGNSPVVYFQ						
	APENFPPLGFMNRQVVLAYANPTLSAV"						
gene	5941..7986						


```

EDVHLMGFTLSGQKVDSPLEASKEWAFRTGVPKKNVYTRGERAKTCYNTSVTDPSG
KSLLDPPNTRIDPKCKTHHQONPHAGQIALHMGCAFLDRIASTMYRGKVP
TEGNTAAMVNTKVMHMLPSQGGIRHNNLTSINKYTWSSNGQNTDTCFQALQY
NSTKNOTCAPSKIPPLPTAPPEIKLITPTDANKLNTDPSDDSDIATSGSGGER
EPIHTTSVATWQGLSTMPPTSPQSPSTPOQGNNTNHSQDAVTELDKNTTAQPSMP
PHNTTTISNTNTHKFNFTSAPLQNTNDNTQITITENEOTSAPISITLPTGNPT
AKSTSSKGGTAPANTNEHTSPPTPSSTAQHLVFRKRSILWREGDMFPLDGL
LNAIDPDPVNTKTIIDESSSSGASAEEDQHASPNISLITLSPNNINENTAYSGEN
ENDCAELRMSVQDDDAAGLSWIPFGGLEGTITAGLKNQNNICEDLRNLANQT
AKSLSELPNTTEERTSLINRHAIDFLITRWGCTKVLGPDCCIGIEDLSKNSIQEI
DQIKDEQKEGTGMLGQKMTSDWGLVNTLNLGILLIALLIALSCICRIFTKYIG"
8869. .9714
/gene="VP30"
8869. .9714
/gene="VP30"
/codon_start=1
/product="VP30 protein"
/db_xref="GI:40388392"
/translation="MQPRGRSRTRHQVPTIYHETQLPSKPHYTNVHPRARMSST
RSSAESPTNHLIPRAPSPSTFNLSKPPPPKDMCRNMKIGLPCADPTCNRDHDLNLT
NRELLLMARMLKNTPTDKTFRSPQCSPLSKGLSKDKQKQDVLTEHLNGLHSY
LHRSIGKLDLSLAALSITCAGIRKTRNSLINTMTLHNNHENLPQDQNGVTKQTY
TGIHLDKGQPEALWQWDKRSISLFOAALYVNNPTPCSSISVQASVDHFILPQS
QCKGO"
10207. .10968
/gene="VP24"
10207. .10968
/gene="VP24"
/codon_start=1
/product="VP24 protein"
/db_xref="GI:40388393"
/translation="MAELSTRYNLPANVTENSINDLNSTARWIKESVPGWTVKWN
FVPHIPTMTLHLHLKSNFVPEWQOTRNFSLFKPNKSTIIEPFLAILLGLAL
KQDELQSLIPGRSIVHMLSEWLVTSNAHISPNIIGVILYLSDFMKILMAGVNF
ENKMTLHVNHDGKPSLEIKLTGQQLITRVNMGFLVVRIDIEFCESGTVLSES
VVGVLVAELRHESQMEKGQPLNLTQMNKIAL"
11481. .18476
/gene="L"
11481. .18476
/gene="L"
/notes="major component of RNA-dependent RNA-polymerase"
/product="L protein"
/protein_id="AA085466.1"
/db_xref="GI:40388394"
/translation="MQHPTQVDPARLSPIILDOCLLARSILGLYSHVSHNPKLNCR
IPHHIYRLNSTALTKFLQNCISILVPHSINWDLITSIQDAINHVDDFKYLLPSEL
VYIANDNEFLKAYLNLKILGDLHVSASARSCBDFSPKNPYMGMLLLVHLSQLAR
RIKQGRSLRNMKFIGTDLFLFIADFVIFKVPVKTIIRNAVLSQASKPLRIWYRD
QNLTPYLCDDPEIVSAVYECFIMIKDVFIERVNTWEICARAWLESDSGADYPLDVL
GELYNQGOI I AMYLEDGFKLKHLEPLCVSCIQTHGFTPRKYWFQSQMIKSYVDEL
HDLNKLQISDNKSCAONFIKTIYQAKLTPOYCELPLOKHGHPVLYNDVALDKV
KHAQSTKILKPKWMEFCVPKFIYAKNHHVSSQSWYKTHDLHLTPYLQHIUSNS
FQSAETIOHLMWEIFVSHBPLFSTKISDLSISIFIKRATAVNOECWDSVFRSIVLY
GKPLVYRVNQVTEALADLAGAKAFPPSNMVMVTEREQEALYAPSRNFSLXKEINLRTF
NGLPYRQKSRPEQIGQADFSLNQILEFAELKYLAPSRNFSLXKEINLRTF
ATVRGASVFTDLKYNLAFRYEFTRHFDYCNRCYGVKNLPDMWFLIPLCYMHVSDP
YSPHCVTEDNENRPPDCANA VHYHLGIGLQKLTWCISCAOITVELTKLPLKLS
CYMGNOQITLISLFPIDAPNDYQNEAELNARVARELIATITGSGILFKPEETFVH
SGFIYFGRKQYLVOLPQSLMARCPGLSDSIDDLQGLSLASIGTSFERGTSETRH
IFPSRWIASFHSMLAINLNQHLGFLGNIIDISCPKPLTFSEKLIATLTPQVLG
LSFLNPEKLYENISDPTSGILFQNLKALEFLEKEELFYILI SKRPLGADASDFVWNP
LGLNVPKSEIITFLROTRENIITTSNRIINSLFHIGSDLEDORVCEWLLSNPMV
SRFADIDSRTPSGKLOVLGYLEGTRTILASRTISLATEGTMKMLRELTRNWKSW
FSYIDALDDLESLEKFTCTVDVANFLRAYSVDVLGKRLIGATLFCALLEQFEKWK
INLSLEDRQFNLSSDSKSTINLLPYDCKELEGNSTELNYVSCALDRKVVQKHS
VNLAWTIGNRNPYIGSRTEDKIGYPPLRVNCPSAALKEAIEWVSRLLWVTQGTADRE
KLLIPLLSNRVNLDTVLNPLFTHYSNIVHRYNDQYQHSFMANRMSNTSTRAIIS
TNTLGKYGAGGQAAIDSNIIIFONTINLGAVLIDIALSLAKLSASNSVTFRLMLNKCCT
RHVPSEYLYFDKPLVDLKNYDNELVNDNPLCSGIKGRIGRVSNSVTSLSLNSVDI
GSYDFPTAAWTLGSETVGSIFDSSESSQSDTDPISSGCTKTFTVTHFLVFPVESYAFG

```

ORIGIN

	Query Match	Best Local Similarity	Score	DB	Length
	Matches	1303	Conservative	0	Mismatches
				2	Indels
				0	Gaps
Qy	1	ATGAAGACCACATGTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAAATCTCCCC	60		
Db	5941	ATGAAGACCACATGTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAAATCTCCCC	6000		
Qy	61	ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	120		
Db	6001	ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	6060		
Qy	121	CTCAGAAGACAGAAAGAGTCCATCTGATGGGATTCACACATGAGTGGGCAAAAATGTCT	180		
Db	6061	CTCAGAAGACAGAAAGAGTCCATCTGATGGGATTCACACATGAGTGGGCAAAAATGTCT	6120		
Qy	181	GATTCCCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACTCTCCCAAGAT	240		
Db	6121	GATTCCCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACTCTCCCAAGAT	6180		
Qy	241	GTTGAGTACACAGAGGGGAGGAGGCAAAACATGCTCAATATAAGTGTAAAGCGATCCC	300		
Db	6181	GTTGAGTACACAGAGGGGAGGAGGCAAAACATGCTCAATATAAGTGTAAAGCGATCCC	6240		
Qy	301	TCTGMAAATCTTGTCTGTAGATCTCTTACCAACATCCGTGACTATCCGAAATGCAAA	360		
Db	6241	TCTGMAAATCTTGTCTGTAGATCTCTTACCAACATCCGTGACTATCCGAAATGCAAA	6300		
Qy	361	ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTATGG	420		
Db	6301	ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTATGG	6360		
Qy	421	GGAGCATTTTTTCTGTATGATCGCATTCCTCCACAAACATGTATCCAGGACAAAGTCTTC	480		
Db	6361	GGAGCATTTTTTCTGTATGATCGCATTCCTCCACAAACATGTATCCAGGACAAAGTCTTC	6420		
Qy	481	ACTGAAGGACACATAGCAGCTATGATGTCTAATAGCAGTGCACAAATGATTTTCTCG	540		
Db	6421	ACTGAAGGACACATAGCAGCTATGATGTCTAATAGCAGTGCACAAATGATTTTCTCG	6480		
Qy	541	CGGCAAGGACAGGATCCGCTCATATGAATCTGACTTCTACTATAAATATTTGGACAAGT	600		
Db	6481	CGGCAAGGACAGGATCCGCTCATATGAATCTGACTTCTACTATAAATATTTGGACAAGT	6540		
Qy	601	AGTAACGGAAACGCAAAACGAATGACATGATGTTTGGCGCTCTTCAAGAATACAAATCT	660		
Db	6541	AGTAACGGAAACGCAAAACGAATGACATGATGTTTGGCGCTCTTCAAGAATACAAATCT	6600		
Qy	661	ACAAGAACCAAAACATGCTCCGCTCCAAATACTCTCACCACTGCCACACCGCGCTCG	720		
Db	6601	ACAAGAACCAAAACATGCTCCGCTCCAAATACTCTCACCACTGCCACACCGCGCTCG	6660		
Qy	721	GAGATCAAACTCACAAGCACCCCAACTGATGCCCAAACTCAATACCACGACCCCAAGC	780		
Db	6661	GAGATCAAACTCACAAGCACCCCAACTGATGCCCAAACTCAATACCACGACCCCAAGC	6720		

QY	781	AGTGATGATGAGGACCTCCCAACATCCGGCTCAGGGTCGGGAGACGAGAACCCCAACACA	840
DB	6721	AGTGATGATGAGGACCTCCCAACATCCGGCTCAGGGTCGGGAGACGAGAACCCCAACACA	6780
QY	841	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA	900
DB	6781	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA	6840
QY	901	CAACCAAGCACGCCACACGCAAGGAGGAAACAACAACAACCATTTCCCAAGATGCTGTGACT	960
DB	6841	CAACCAAGCACGCCACACGCAAGGAGGAAACAACAACAACCATTTCCCAAGATGCTGTGACT	6900
QY	961	GAACTAGACAAATAAACAACAACGTCGACAAACCGTCCATGCCCCCTCATTAACACTACCA	1020
DB	6901	GAACTTGACAAATAAACAACAACGTCGACAAACCGTCCATGCCCCCTCATTAACACTACCA	6960
QY	1021	ATCTCTACTAAACAACCTCCAAACAACACTTTCAGCACTCTCTGCAACCATTAACAAC	1080
DB	6961	ATCTCTACTAAACAACCTCCAAACAACACTTTCAGCACTCTCTGCAACCATTAACAAC	7020
QY	1081	ACCACCAATGACAAACAACAGGACCAATCACTGAAAATGAGCAAAACCAAGTGCCTCCG	1140
DB	7021	ACCACCAATGACAAACAACAGGACCAATCACTGAAAATGAGCAAAACCAAGTGCCTCCG	7080
QY	1141	ATAACAACCTGCTCCCAACGGGAATCCCAACACAGCAAGCAAGAGCAGCAGCAAAAA	1200
DB	7081	ATAACAACCTGCTCCCAACGGGAATCCCAACACAGCAAGCAAGAGCAGCAGCAAAAA	7140
QY	1201	GGCCCCGACACAGCGCACCAACAGCAACATGAGCAATTCACCAAGTCCCTCCGCCACC	1260
DB	7141	GGCCCCGACACAGCGCACCAACAGCAACATGAGCAATTCACCAAGTCCCTCCGCCACC	7200
QY	1261	CCAGCTCGACTGCAACAACATCTTGATATATTTTCAAGAAAGCGA	1305
DB	7201	CCAGCTCGACTGCAACAACATCTTGATATATTTTCAAGAAAGCGA	7245
RESULT 7			
AX717721			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1. .2051			
/organism="Marburg virus (strain Musoke)"			
/mol_type="unassigned DNA"			
/db_xref="taxon:3372"			
/note="chimeric molecule between Marburg virus			
Glycoprotein 1 and Marburg virus Glycoprotein 2"			
ORIGIN			
Query Match			
Best Local Similarity 99.1%; Pred. No. 0;			
Matches 1296; Conservative 0; Mismatches 0; Indels 12; Gaps 1;			
QY	1	ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGGGACAAAAATCTCCCC	60
DB	10	ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGGGACAAAAATCTCCCC	69
QY	61	ATTTTAGATAGCTAGTAAATCAACCCCAAAATGCGATTCGGTATGCTCCGGA	120
DB	70	ATTTTAGATAGCTAGTAAATCAACCCCAAAATGCGATTCGGTATGCTCCGGA	129

QY	121	CTCAGAAAGACAGAGACGTCATCTGTATGGGATTCACACTGAGTGGGCAAAAAAGTTGCT	180
DB	130	CTCAGAAAGACAGAGACGTCATCTGTATGGGATTCACACTGAGTGGGCAAAAAAGTTGCT	189
QY	181	GATTCCCTTTTGGAGGCATCCAAAGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAA	240
DB	190	GATTCCCTTTTGGAGGCATCCAAAGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAA	249
QY	241	GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATATAGTGTAAACGGATCCC	300
DB	250	GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATATAGTGTAAACGGATCCC	309
QY	301	TCTGAAATCTTCTGCTTAGATCTCTCTACCAACATCCGTGACTATCCGAAATGCAAA	360
DB	310	TCTGAAATCTTCTGCTTAGATCTCTCTACCAACATCCGTGACTATCCGAAATGCAAA	369
QY	361	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGGATCGCCTTCATTTATGG	420
DB	370	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGGATCGCCTTCATTTATGG	429
QY	421	GGAGCATTTTCTGTATGATCGCATTCCTCCACAACATGTACCGAGGCAAGCTCTTC	480
DB	430	GGAGCATTTTCTGTATGATCGCATTT-----ATGTACCGAGGCAAGCTCTTC	477
QY	481	ACTGAAGGAAACATAGCAGCTATGATGTCAATAGACAGATGCAACAAATGATTTTCG	540
DB	478	ACTGAAGGAAACATAGCAGCTATGATGTCAATAGACAGATGCAACAAATGATTTTCG	537
QY	541	CGGACAGACAGGGTACCGTCATATGAATCTGACTTCTACTATAAATATTTGGACAGT	600
DB	538	CGGACAGACAGGGTACCGTCATATGAATCTGACTTCTACTATAAATATTTGGACAGT	597
QY	601	AGTAACGGAACGCAAAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAA	660
DB	598	AGTAACGGAACGCAAAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAA	657
QY	661	ACAAAGAACCAAAACATGTCTCCGTCCTCCAAATACTCTCCACCACTCCCAAGCCGCTCG	720
DB	658	ACAAAGAACCAAAACATGTCTCCGTCCTCCAAATACTCTCCACCACTCCCAAGCCGCTCG	717
QY	721	GAGTCAAACTCACAGCACCCCAACTGATGGCACCACAACTCATATACCGACCCCAAGC	780
DB	718	GAGTCAAACTCACAGCACCCCAACTGATGGCACCACAACTCATATACCGACCCCAAGC	777
QY	781	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCACACA	840
DB	778	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCACACA	837
QY	841	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA	900
DB	838	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA	897
QY	901	CAACCAAGCACGCCACACAGAGGAGGAAACAACAACCATTTCCCAAGATGCTGTGACT	960
DB	898	CAACCAAGCACGCCACACAGAGGAGGAAACAACAACCATTTCCCAAGATGCTGTGACT	957
QY	961	GAACTAGACAAATAAACAACAACGTCGACAAACCGTCCATGCCCCCTCATAACTACCA	1020
DB	958	GAACTAGACAAATAAACAACAACGTCGACAAACCGTCCATGCCCCCTCATAACTACCA	1017
QY	1021	ATCTCTACTAAACAACCTCCAAACAACCTTCAGCACTCTCTCTGCAACCATTAACAAC	1080
DB	1018	ATCTCTACTAAACAACCTCCAAACAACCTTCAGCACTCTCTCTGCAACCATTAACAAC	1077
QY	1081	ACCACCAATGACACACAGAGCAACATCACTGAAATGAGCAAAACCAAGTGCCTCCG	1140
DB	1078	ACCACCAATGACACACAGAGCAACATCACTGAAATGAGCAAAACCAAGTGCCTCCG	1137
QY	1141	ATAACAACCTGCTCCCAACGGGAATCCCAACAGCAAGAGAGCAGCAGCAAAAA	1200
DB	1138	ATAACAACCTGCTCCCAACGGGAATCCCAACAGCAAGAGAGCAGCAGCAAAAA	1197

```
QY 1201 GGGCCGCGCACAAACGGCACCAACAGCAAAATGAGCATTTTACCAGTCTCTCCCCCACC 1260
Db 1198 GGGCCGCGCACAAACGGCACCAACAGCAAAATGAGCATTTTACCAGTCTCTCCCCCACC 1257
QY 1261 CCAGCTCGACTGCACAACTCTTGATATTTTTCAGAGAAAGCGATCG 1308
Db 1258 CCAGCTCGACTGCACAACTCTTGATATTTTTCAGAGAAAGCGATCG 1305

RESULT 8
AF005735      2046 bp      RNA      linear      VRL 23-JAN-1998
LOCUS      Marburg virus strain M/Germany/Marburg/1967/Ratayczak glycoprotein
DEFINITION      precursor (GP) gene, complete cds.
ACCESSION      AF005735
VERSION      1 (bases 1 to 2046)
KEYWORDS      Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and
SOURCE      Feldmann,H.
ORGANISM      Variation in the glycoprotein and VP35 genes of Marburg virus
strains
JOURNAL      Virology 240 (1), 138-146 (1998)
MEDLINE      98110148
PUBMED      9448698
REFERENCE      2 (bases 1 to 2046)
AUTHORS      Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
TITLE      Direct Submission
JOURNAL      Submitted (29-MAY-1997) Special Pathogens Branch, Centers for
Disease Control and Prevention, 1600 Clifton Road, Building 15,
Room SB611, Mail Stop G14, Atlanta, GA 30333, USA

FEATURES
     source
     1..2046
     /organism="Lake Victoria marburgvirus"
     /mol_type="genomic RNA"
     /strain="M/Germany/Marburg/1967/Ratayczak"
     /db_xref="taxon:11269"
     1..2046
     /gene="GP"
     1..2046
     /gene="GP"
     /note="virion surface glycoprotein"
     /codon_start=1
     /product="glycoprotein precursor"
     /protein_id="AAC40460.1"
     /db_xref="GI:2459880"
     /translation="MKTCLFLSLILIQIKTLPILEIASNNQPNVDVSCSGTLQK
     EDVHLMGFTLSGQKVDSPLEASKEWAFRTGVPKKNVEVEGEAKTCVNI SVTPDSG
     KSLLDPPNTNIRDYPKCKTIHI OGNPHAOGIALHMGAPFLYDRIASTTMYGRVF
     TEGNTAAMLVNKHVHMIFESROGQYRHNLSTNKYTSNNGTQTDNDCGFLQLEY
     NSTKQTCAPSKIPPLPAPRIEPTPTPQEGNNDHSQGTVPKNTNTAQPSP
     EPTTSDAVTKQGLSTWPPSPQFSTPQEGNNDHSQGTVPKNTNTAQPSP
     PHNTAISNTSKNFTLSLQNTNYDTSATENEQTSAPSKTTLPTGLNLT
     AKSTNTKPTNTIAPNTMNGHLTSPPTNPPTQHLVYFRKRSILWREGDMFPLDG
     LNPADIDPVPNTKIIPDESSSGASAEODHASPNI SLTSLYFPNIENMTAVSGEN
     ENDCAEALRINSVQEDDLAAGLSWIPFGFGLGYTAGLTKQNNLVCRNLRIANQT
     AKSLELLURVTEFTSLINRHAIDFLLRWGTCKVIGPDCCTCIEDLSRNTSEQI
     DOI KXDEQKRGWGGLGKWWTSWGLVTLNLGILLLSIAVLIALSCIRIPTYIG"
     1..81
     /gene="GP"
     82..2043
     /gene="GP"
     /product="glycoprotein"
     1930..2019
     /gene="GP"
     /note="encodes transmembrane anchor peptide"

sig_peptide
mat_peptide
misc_feature
misc_feature

Query Match      62.5%; Score 1151.4; DB 14; Length 2046;
ORIGIN
```

```
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1209; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATCAAGGGACAAAAAATCTCCCC 60
Db 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATCAAGGGATAAAAAATCTCCCT 60
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
QY 121 CTCAGAGACAGAGAGCGTCCATCTGATGGATTTCACTAGAGTGGGCAAAAAGTGTCT 180
Db 121 CTCAGAGACAGAGAGATGTCATCTGATGGATTTCACTAGAGTGGGCAAAAAGTGTCT 180
QY 181 GATTCCCTTTTGGAGGCATCCAAGCGATGGGCTTTTCAGACAGAGGTGTAACCTCCCAAGAT 240
Db 181 GATTCCCTTTTGGAGGCATCCAAGCGATGGGCTTTTCAGACAGAGGTGTAACCTCCCAAGAT 240
QY 241 GTTGAGTACAGAGGGGGAGAGCCAAAACATGCTACAATAAAGTGTAAAGATATCC 300
Db 241 GTTGAGTATACAGAGGGGGAGAGCCAAAACATGCTACAATAAAGTGTAAAGATATCC 300
QY 301 TCTGGAATCTTCTGTTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Db 301 TCTGGAATCTTCTGTTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
QY 421 GGAGATTTTTTCTGATGATGCGATTCCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
Db 421 GGAGATTTTTTCTGATGATGCGATTCCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
QY 481 ACTGAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTCTCG 540
Db 481 ACTGAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTCTCG 540
QY 541 CGGCAAGCAACAGGGTACCGTCAATATGATCTGACTTCTACTATAATAATTTTGGACAAAT 600
Db 541 AGGCAAGCAACAGGGTACCGTCAATATGATCTGACTTCTACTATAATAATTTTGGACAAAT 600
QY 601 AGTAACGAAACGCAAAACGAATGACACTGGATGTTTGGCGCTCTTCAAGAATATCAATTTCT 660
Db 601 AACCAAGCAACAGGGTACCGTCAATATGATCTGACTTCTACTATAATAATTTTGGACAAAT 660
QY 661 ACAGAGAACCAACATGCTCGTCCGTCCAAAATACCTCCACCACTGCCCACAGCCGCTCCG 720
Db 661 ACAGAGAACCAACATGCTCGTCCGTCCAAAATACCTCCACCACTGCCCACAGCCGCTCCG 720
QY 721 GAGATCAAACTCACAGACACCCCAACTGATGTCACCAAACTCAATACCAAGGACCAAGC 780
Db 721 GAGATCAAACTCACAGACACCCCAACTGATGTCACCAAACTCAATACCAAGGACCAAGC 780
QY 781 AGTGATGATGAGGACCTCGCAACATCCCGCTCAGGGTCCGAGAAACGAGAACCCCAACA 840
Db 781 AATGATGATGAGGACCTCATAACTCCGCTTCCGAGTCCGAGAAACGAGAACCCCTATACA 840
QY 841 ACTTCTGATGGGTACCAAGCAAGGGCTTTTCATCAATGCGCACCACTCCCTCCCTACCA 900
Db 841 ACTTCTGATGGGTACCAAGCAAGGGCTTTTCATCAATGCGCACCACTCCCTCCCTACCA 900
QY 901 CAACCAAGCAGCCACAGCAAGGAGAAACAAACCAAACTTCCCAAGATGCTGTGACT 960
Db 901 CAACCAAGCAGCCACAGCAAGGAGAAACAAACCAAGATGCTGTGACT 960
QY 961 GAACTAGACAAAAATAACAACTGACAAACCGTTCATGCCCTCCCTCATAACCTACCA 1020
Db 961 GAACTAGACAAAAATAACAACTGACAAACCGTTCATGCCCTCCCTCATAACCTACCA 1020
QY 1021 ATCTCTACTAACCAACCTCCCAACCAACTTCAGACTCTCTCTGACCACTTACCAAAAC 1080
Db 1021 ATCTCTACTAACCAACCTCCCAACCAACTTCAGACTCTCTCTGACCACTTACCAAAAC 1080
```

```
Db 1021 ATCTCTACTAACACACCTCCCAAGAACAACTTCAGACCCCTCTCTGTATCATCAACAAAC 1080
QY 1081 ACCACCAATGACACACACAGACGACAACTCACTGAAATGACAAACAGTCCGCCCTCG 1140
Db 1081 ACCACCAATGACACACACAGACGACGACCACTGAAATGACAAACAGTCCGCCCTCG 1140
QY 1141 ATAACAACCTCTCCCTCAACGCGGAATCCACACAGCAAGAGCACCAGCAGCAAAAA 1200
Db 1141 AAAACAACCTCTCCCTCAACGCGGAATCTTACACAGCAAGAGCACAACACAGAAA 1200
QY 1201 GCGCCGCGACACAGCGCACCAACAGCAAAATGAGCATTTCCAGTCTCCCTCCACC 1260
Db 1201 GCGCCGCGACACAGCGCACCAATATGACAAATGGGCAATTAACAGTCTCCCTCCACC 1260
QY 1261 CCCAGCTCGACCTGCACCAACATCTTGATATTTTTCAGAAAGAGCGA 1305
Db 1261 CCCAAGCGACACACACACATCTTGATATTTTCAGAAAGAGCGA 1305

RESULT 9
LOCUS MAVSPA 2948 bp RNA linear VPL 22-JAN-1996
DEFINITION Marburg virus genomic RNA of GP gene.
ACCESSION X68493
VERSION X68493.1 GI:296960
KEYWORDS gp gene; structural protein.
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 2948)
AUTHORS Bukreyev,A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.
TITLE The GP-protein of Marburg virus contains the region similar to the
'immunosuppressive domain' of oncogenic retrovirus P15S proteins
FEBS Lett. 323 (1-2), 183-187 (1993)
JOURNAL 93265932
MEDLINE 8495737
PUBMED
REFERENCE 2 (bases 1 to 2948)
AUTHORS Bukreyev,A.A., Volchkov,V.E., Blinov,V.M., Dryga,S.A. and
Netesov,S.V.
TITLE The complete nucleotide sequence of the Popp (1967) strain of
Marburg virus: a comparison with the Musoke (1980) strain
Arch. Virol. 140 (9), 1589-1600 (1995)
JOURNAL 96028047
MEDLINE 7487490
PUBMED
REFERENCE 3 (bases 1 to 2948)
AUTHORS Bukreyev,A.A.
TITLE Direct Submission
Submitted (28-SEP-1992) A.A. Bukreyev, All-Union Inst. of Molecular
Biology, NPO 'Vector', Koltsovo, Novosibirsk region, 633159, USSR
Biological Location/Qualifiers
1..2948
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="Popp"
/db_xref="taxon:11269"
6..2851
/gene="GP"
(6.17)..(2841.2851)
/gene="GP"
122..2167
/gene="GP"
/codon_start=1
/product="structural protein"
/protein_id="CAA48507.1"
/db_xref="GI:296961"
/db_xref="GOA:P32524"
/db_xref="UniProt/Swiss-Prot:P32524"
translation="WKTCLFSLILIQIKLPILEIASNNPONVDVSCSGTLQKT
EDVHLMGFTLSQKVDSPLEASKWAFETGVPKKNVEVTEGEAKTCYNI SVTPDSG
KSLILDPTNIDYPKCKTIHHIQONPHAQIALHMGAFELYDRIASTMYGRVF
TSGNTAAMIVNKTVMHMIIFSRQGGYRHNULTSTNKYMTSSNGTQNDTDCFGALQGY
NETQKQTCAPSKIPSLPTARPEIKPTSTPTDATTTLNTDTPNDDDELITSGSGSEQ
```

ORIGIN

```
Query Match 62.4%; Score 1148.2; DB 14; Length 2948;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTTCCTTATCAGTCTTATCTTAATTCAGGAGCAAAAAATCTCCCC 60
Db 122 ATGAAGACCAATGCTCTTATCAGTCTTATCTTAATCCAAGGATATAAAATCTCCCT 181
QY 61 ATTTTAGAGATAGTAGTAATAACCAACCCCAAAATGTGGATTGGTATGTCCTCCGAAT 120
Db 182 ATTTTAGAGATAGTAGTAATAACCAACCCCAAAATGTGGATTGGTATGTCCTCCGAAT 241
QY 121 CTCCAGAGACAGAGAGCGTCCATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 242 CTCCAGAGACAGAGAGATGTCCATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 301
QY 181 GATTCCCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGAGTGTACCTCCCAAGAT 240
Db 302 GATTCCCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGAGTGTACCTCCCAAGAT 361
QY 241 GTTAGATACACAGAGGGGAGAGCAAAACATGCTACATATAAGTGTAAAGTAAAGTATCC 300
Db 362 GTTAGATACACAGAGGGGAGAGCAAAACATGCTACATATAAGTGTAAAGTAAAGTATCC 421
QY 301 TCTGGAATCTCTGCTGTAGATCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Db 422 TCTGGAATCTCTGCTGTAGATCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 481
QY 361 ACTATCCATCATATTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db 482 ACTATCCATCATATTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 541
QY 421 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACAAATGTACCGAGGCAAGTCTTC 480
Db 542 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACAAATGTACCGAGGCAAGTCTTC 601
QY 481 ACTGAAGGGAACATAGCAGCTATGATGTCATAGACAGTGCACAAATGATTTCTCG 540
Db 602 ACTGAAGGGAACATAGCAGCTATGATGTCATAGACAGTGCACAAATGATTTCTCG 661
QY 541 CGGCAAGGACAGGATACCGTACATATGAATCTGACTTCTACTATAAATATTGGCAAGT 600
Db 662 AGGCAAGGACAGGATACCGTACATATGAATCTGACTTCTACTATAAATATTGGCAAGT 721
QY 601 AGTAACGGAACCAAAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGATATCAATTT 660
Db 722 AACATGGAACACAAACGAATGACACTGGATGTTTCGGTGTCTTCAAGATATCAATCTC 781
QY 661 ACAAGAACCAAAACATGTCCTCCGTCACAAATACCTCCACCACTGCCACAGCCCGTCCG 720
Db 782 ACGAAGAAATCAAAACATGTCCTCCGTCACAAATACCTCCACCACTGCCACAGCCCGTCC 841
QY 721 GAGATCAAACTCAAAAGCAACCCCAACTGATGCACCAAACTCAATACCAAGGACCCCAAG 780
Db 842 GAGATCAAACTCAAAAGCAACCCCAACTGATGCACCACTCAATACCAAGGACCCCAAG 901
QY 781 AGTATGATGAGGACCTCGCAACATTCGGCTCAGGGTCCGAGAAACGAGAAACCCACACA 840
Db 902 AATGATGATGAGGACCTCATAACTCCGTTTCAGGGTCCGAGAAACGAGAAACCCCTACA 961
QY 841 ACTTCTGATGGGTCCACAAAGCAAGGGTTCATCAAAATGCCACCACTCCCTCAACA 900
Db 962 ACTTCTGATGGGTCCACAAAGCAAGGGTTCATCAAAATGCCACCACTCCCTCAACA 1021
```

QY 901 CAAACGACGCCACGACGAGGAGAAACACACAAACCAATTCACAGATGCTGTGACT 960
Db 1022 CAACCAAGCACCCACAGCAAGAGAAACACACAGACCAATTCACAGGTACTGTGACT 1081
QY 961 GAACTAGACAAAATAACACAACTGCACAAACCGTCCATGCCCTCCATACAACTACACACA 1020
Db 1082 GAACCCACAAACCAACACACAGGACACACGTCCTATGCCCTCCACACACCACTGCA 1141
QY 1021 ATCTCTACTAACACACCTCTCAAAACACAACTTCAGCATCTCTCTGACCACTTACAAAC 1080
Db 1142 ATCTCTACTAACACACCTCTCAAAACAACTTCAGCATCTCTCTGATCACTACAAAC 1201
QY 1081 ACCCAATGACACACACAGGACCAATCACTGGAATGACAAACAGTGCCTCCCTCG 1140
Db 1202 ACCCAATATGACACACAGGACGACCACTGGAATGACAAACAGTGCCTCCCTCG 1261
QY 1141 ATAAACAACCTGCTCCCAACGGGAATCCACACAGCAAGAGACACGACGACGAAAAA 1200
Db 1262 AAAACAACCTGCTCCCAACAGGAATCTTACACAGCAAGAGACCACTAACACAGAAA 1321
QY 1201 GGGCCGCGACAAACGGCACCAACACGACAAATGAGCATTTTCCAGTCTCTCCCTCCACC 1260
Db 1322 GGGCCGCGACAAACGGCACCAATATGACAAATGGCATTTTAAACAGTCTCTCCCTCCACC 1381
QY 1261 CCCAGCTCGACGACCAACATCTTGTATATTTTCAGAGAAACGGA 1305
Db 1382 CCCAAGCCGACACACAAATCTTGTATATTTTCAGAGAAACGGA 1426

RESULT 10
MVIRPR Marburg virus (Popp) NP, VP35, VP40, GP, VP30, VP24, L genes. linear VRL 22-JAN-1996
LOCUS 229337
DEFINITION 229337.1 GI:450908
ACCESSION
VERSION
KEYWORDS GP protein; L protein; NP protein; NP protein; vp24 protein; vp30 protein; vp35 protein; vp40 protein.
SOURCE Lake Victoria marburgvirus
ORGANISM Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 2851 to 5818)
AUTHORS Bukreyev,A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.
TITLE The VP35 and VP40 proteins of filoviruses. Homology between Marburg and Ebola viruses
JOURNAL FEBS Lett. 322 (1), 41-46 (1993)
MEDLINE 93245956
PUBMED 8482365
REFERENCE 2 (bases 5824 to 8669)
AUTHORS Bukreyev,A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.
TITLE The GP-protein of Marburg virus contains the region similar to the 'immunosuppressive domain' of oncogenic retrovirus P15E proteins
JOURNAL FEBS Lett. 323 (1-2), 183-187 (1993)
MEDLINE 93265932
PUBMED 8495737
REFERENCE 3 (bases 1 to 19112)
AUTHORS Bukreyev,A., Volchkov,V.E., Blinov,V.M., Dryga,S.A. and Netesov,S.V.
TITLE The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain
JOURNAL Arch. Virol. 140 (9), 1589-1600 (1995)
MEDLINE 96028047
PUBMED 7487490
REFERENCE 4 (bases 1 to 19112)
AUTHORS Bukreyev,A.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-1994) Bukreyev A.A., Institute of Molecular Biology, Laboratory of molecular virology, Koltsovo, Novosibirsk region, Russia, 633159
FEATURES Location/Qualifiers
source 1..19112
organism="Lake Victoria marburgvirus"
mol_type="genomic RNA"

gene
CDS
/strain="Popp"
/db_xref="taxon:11269"
103..2190
/gene="NP"
103..2190
/gene="NP"
/function="nucleoprotein"
/citation=[3]
/number=1
/codon_start=1
/label=NP
/product="NP protein"
/protein_id="CAA82536.1"
/db_xref="GI:450909"
/db_xref="GOA:P35263"
/db_xref="UniProt/Swiss-Prot:P35263"
/translation="MDLHLELGTPTAPHRNKVILFDTNHQSVCNQIIDAINS
GIDGLLEGGLTLCVEHYNSDKFNTSPIAKYLRDAGVDFDVKNAADATFPLDV
IPNPHYSPLIALKLTLESTESQGRIGLFLSCLEPLKLVGDRASIEKALQVTV
GQTFPSGLLIVKTVLEFILQKDSGVTLPVLRVTSKVNEVASFKQALSNLARHGEYA
PFARVLNLSGNNLEHGLYPOLSAIALGVATAHGSTLAGVNVGQYQQLREAAHDAE
KLRHHEHOIAIAEDDEERKILEQPHLOKTEITHSOTLAVLSOKREKLAAAEIE
NNIVEDGFKQSONQVSFLNDPTPVEVTVOARPINRPTALPPVDNKKIEHSTEDS
SSSSFVDLNDPFPALLNEDEDTLDSVMIPTSTTSREFOGIPAPRQSQDLNNSQKQE
DESTNPIKKFLRYQELPPVQEDDESEYTTDSQSDIDQSGDNQGVLDLPPPLIYQE
KRODPIQHPAVSSQDPFGSGIDVNGDILEPIRSFSPSQEDTRAREAYELSPDFTN
YEDNQNPQORVTVTKGRFTFLYPNDLLQTNPPESLITALVEEQNPVSAKELQADWPD
MSFDRRHVAMNL"
2944..3933
/gene="VP35"
2944..3933
/gene="VP35"
/function="minor component of RNA-dependent
RNA-polymerase?"
/citation=[1]
/citation=[3]
/number=2
/codon_start=1
/label=VP35
/product="VP35 protein"
/protein_id="CAA82537.1"
/db_xref="GI:450910"
/db_xref="UniProt/Swiss-Prot:Q03039"
/translation="MDDSSYMQVSEGLMTGKVPIDQVFGANPSEKLRKRPKGTG
LQCSPLMSKATSDTDIVMDLIIVKTKLADLLIPINRQISDIQSTLNVEITRVHEIER
OLHEITPVLMGRGLEAISKGMSEMLAKYDHLVISTGTTPAPAAAFAYLNEHGVPFP
QPAIFKDLGVAQAQCSKGTWVNETTDAADKMSKVLSESETEKPNLSAKDLALLF
THLFGNTPPHILAQVLSKIAYKSGKGAFLDAFHQILSEGENAQAALTRLSPTDAP
LGVVPPVIRVKNFQTPRCPQKSLRAVPPNPITDKGWVCVYSSSQGETRALKI"
4567..5478
/gene="VP40"
4567..5478
/gene="VP40"
/function="matrix protein?"
/citation=[1]
/citation=[3]
/number=3
/codon_start=1
/label=VP40
/product="VP40 protein"
/protein_id="CAA82538.1"
/db_xref="GI:450911"
/db_xref="GOA:Q03040"
/db_xref="UniProt/Swiss-Prot:Q03040"
/translation="MASSNYNTYQVLPVPPYADHGANQLIPADLSNQOQITPNVY
GDLNDDQKGNVCHAFLEAIIIDISATNPTVKGVPKAWFLGINSFEYFLAHTVAA
LLTCSYITQTFHNGQKFRVIRLIGTGIPAPHLRMLREGNQAFIQNMVYPRNFSTNQ
TYNTNLVLSVQKLPDDAWRSPDKLTIGNTMHPAVSIHPLPFIPLVPTVRKQAYRQHK
NPNNGPLIAISGILHQLRVKVPKTSFLRI SLPADMFPSVKEGMMKRGENSEPVVYFQ
APENFPLNGENNRQVVLAYANPTLSAV"
5940..7985
/gene="GP"

CDS

```
5940. .7985
/function="GP"
/function="envelope transmembrane glycoprotein"
/citation=[2]
/citation=[3]
/number=4
/codon_start=1
/label=GP
/product="GP protein"
/protein_id="CAA82539.1"
/db_xref="GI:450912"
/db_xref="GOA:P35254"
/db_xref="UniProt/Swiss-Prot:P35254"
/translation="MKTCLFSLILIOGIKTLPLIEASNNQPNQVDSVCSGTLQKT
EDVHLMGFSLQKQVADSLKQVARTGVPKNVEYTEGEAEKTCVNTSVDPDG
KSLLDPPNTIRDYKCKTIRHI QGONPHAOQIALHLCAPLYDRIASITMYGRVRF
TEGNTAMIVKNTVHKMIFSRQOGYRHMNLSTNKYWTSSNGTGTNTDTCFQALQEQ
NSTKQTCAPSKIPFLPTARPEIRKPTSTPTDATTILNTDNNDEDLITSGSGSEQ
EPTTSDAVTKGLSSTWPPSPQSPSTPQEGNNTDHSQGTVPBNKINTTAQSPMP
PHNTTASINTNSKNFSTLSVLQNTYNTDQSTATENEQTSAPSKTTLPTGLNLT
AKSTNNTKGTPTAPNMTNGLTSPSPNPTTQHLVYFRKRSILWREGDMFPLDG
LINAIDPDVPNTKTIIDESSSSGASAEEDOHASPNISLTLSPFNINENTAYSGEN
ENDCAELRIMSVQEDDLAAGLSWIFPFGPIEGLYTAGLTKQNQLVCRLLRANQT
AKSLLURVITEERTFSLNRHADFLLTRWGCTCKVLGPPDCCIGISDLGRNISEQI
DOIKKDEQKEGTGWLGGKWTSDMGVLTNLGILLLSIAVLIALSCICRIPTYIG"
8867. .9712
/genes="VP30"
8867. .9712
/genes="VP30"
/citation=[3]
/number=5
/codon_start=1
/label=VP30
/product="VP30 protein"
/protein_id="CAA82540.1"
/db_xref="GI:450913"
/db_xref="GOA:P41326"
/db_xref="UniProt/Swiss-Prot:P41326"
/translation="WQPRGRSRTRNHOTASSIYHETOLPSKPHYNNHPRARMSST
RSSASSPNHPIPRARPPTNLSPKPPPKDMCKNMKIGLPTDPTCNRDHLDNLIT
NRELLLMARKMLPNTDKTFRSLQDCGSPSLSKGLSKQEQTKDVLFIENLGHILNY
LHRSDIGKDETSLSRAAUSLTGAGIRKTRNSLINTMTLHNLHNLPOQNGVIKQTY
TGIHLDKGGQFAALWQGWDRKRSISLFVQAALYVMNNIPCESSTSVQASYPHFILPQS
QSKGQ"
```

gene

CDS

```
10205. .10966
/genes="VP24"
10205. .10966
/genes="VP24"
/citation=[3]
/number=6
/codon_start=1
/label=VP24
/product="VP24 protein"
/protein_id="CAA82541.1"
/db_xref="GI:450914"
/db_xref="GOA:P41326"
/db_xref="UniProt/Swiss-Prot:P41326"
/translation="WQPRGRSRTRNHOTASSIYHETOLPSKPHYNNHPRARMSST
RSSASSPNHPIPRARPPTNLSPKPPPKDMCKNMKIGLPTDPTCNRDHLDNLIT
NRELLLMARKMLPNTDKTFRSLQDCGSPSLSKGLSKQEQTKDVLFIENLGHILNY
LHRSDIGKDETSLSRAAUSLTGAGIRKTRNSLINTMTLHNLHNLPOQNGVIKQTY
TGIHLDKGGQFAALWQGWDRKRSISLFVQAALYVMNNIPCESSTSVQASYPHFILPQS
QSKGQ"
```

gene

CDS

```
10205. .10966
/genes="VP24"
10205. .10966
/genes="VP24"
/citation=[3]
/number=6
/codon_start=1
/label=VP24
/product="VP24 protein"
/protein_id="CAA82541.1"
/db_xref="GI:450914"
/db_xref="GOA:P41326"
/db_xref="UniProt/Swiss-Prot:P41326"
/translation="WQPRGRSRTRNHOTASSIYHETOLPSKPHYNNHPRARMSST
RSSASSPNHPIPRARPPTNLSPKPPPKDMCKNMKIGLPTDPTCNRDHLDNLIT
NRELLLMARKMLPNTDKTFRSLQDCGSPSLSKGLSKQEQTKDVLFIENLGHILNY
LHRSDIGKDETSLSRAAUSLTGAGIRKTRNSLINTMTLHNLHNLPOQNGVIKQTY
TGIHLDKGGQFAALWQGWDRKRSISLFVQAALYVMNNIPCESSTSVQASYPHFILPQS
QSKGQ"
```

gene

CDS

```
11479. .18474
/genes="L"
11479. .18474
/genes="L"
/function="major component of RNA-dependent
RNA-polymerase"
/citation=[3]
/number=7
/codon_start=1
/label=L
/product="L protein"
/protein_id="CAA82542.1"
```

```
/db_xref="GI:450915"
/db_xref="GOA:P35262"
/db_xref="UniProt/Swiss-Prot:P35262"
/translation="WQHTQYDARLSPIILDQCDLLARSILGLYSHYSHNPKLNR
IPHYILRLNSTALKTFLQNCISILTVPHSHTLHLSIOYDAINHVDFFKYLPSSEL
VKYANWNEFLKYLNLKILGLDHPAPASARQWEDFSKPENPYWGMILLVHLSQLAR
RIKQGRSLRSNWKPIGTDLELFGIADFIIFKVPVKTIIRNAVLSQASQGLRVYWRD
QNLTPYLCDDDEFIVSVASYECFIMIKDVFIERYNTEICARAMELSDGADYLPDLVL
GELYNQGDQIIMYLEDGFKLHKLEPLCVSCIOTHGFTFKPKYWFQQRLESYVEEL
CSLWKKFKISGNKAECAQNFITKIIQGLKTPQOYCELPFSLOKHGHPVLYIDVALDKV
KHAQSVKILPKWMPFPCVPKPIVAKNHYHSGSWYKTTMDLHLTYLRLQHIYSNS
FPSQAEYQHLWEWIFVEHEPLFSTISLISLFIKIRATAVNQBCWDSVDFRSLVIGY
```

Query Match

Best Local Similarity 92.5%; Score 1148.2; DB 14; Length 19112;
Matches 1207; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY	1	ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAAATCTCCCC	60
DB	5940	ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAAATCTCCCC	5999
QY	61	ATTTTAGAGATAGCTAGTAATAACCCCAAAATGCGATTTCGGTATGCTCCGGAAT	120
DB	6000	ATTTTAGAGATAGCTAGTAATAACCCCAAAATGCGATTTCGGTATGCTCCGGAAT	6059
QY	121	CTCAGAGACAGAGACGCTCCATCTGATGGGATTACACATGAGTGGGCAAAAAGTTGCT	180
DB	6060	CTCAGAGACAGAGAGATGTCATCTGATGGGATTACACATGAGTGGGCAAAAAGTTGCT	6119
QY	181	GATTCCCTTTGGAGGACATCAAGGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT	240
DB	6120	GATTCCCTTTGGAGGACATCAAGGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT	6179
QY	241	GTCAGTACACAGAGGGGGAGAGCCAAACATGCTACAATATAAGTGTACGATCC	300
DB	6180	GTTGAGTATACAGAGGGGGAGAGCCAAACATGCTACAATATAAGTGTACGATCC	6239
QY	301	TCTGAAAATCCTTGCTTTAGATGCTCTTACCACATCCGCTGACTATCCGAAATGCAAA	360
DB	6240	TCTGAAAATCCTTGCTTTAGATGCTCTTACCACATCCGCTGACTATCCGAAATGCAAA	6299
QY	361	ACTATCCATCATATTCAGAGTCAAAACCTCATGACAGGGGATCGCTTCATTATGG	420
DB	6300	ACTATCCATCATATTCAGAGTCAAAACCTCATGACAGGGGATCGCTTCATTATGG	6359
QY	421	GGAGCATTTTTCTGTATGATCGCATTCGCTCCCAACAATGTACCGAGGCAAGTCTTC	480
DB	6360	GGAGCATTTTTCTGTATGATCGCATTCGCTCCCAACAATGTACCGAGGCAAGTCTTC	6419
QY	481	ACTGAAGGGAACATAGCAGCTATGATTGTCATTAAGACAGATGCACAAAATGATTTCTCG	540
DB	6420	ACTGAAGGGAACATAGCAGCTATGATTGTCATTAAGACAGATGCACAAAATGATTTCTCG	6479
QY	541	CGGCAAGGACAAAGGTACCGTACATATGAACTGTACTTCTACTATAATATTTGGACAAGT	600
DB	6480	AGGCAAGGACAAAGGTACCGTACATATGAACTGTACTTCTACTATAATATTTGGACAAGT	6539
QY	601	AGTAACGGAACGCAAAACGAATGACATCGATGTTTCGGGCTCTTCAAGAATACAAATCT	660
DB	6540	AACATGGAACACAAACGAATGACATCGATGTTTCGGGCTCTTCAAGAATACAAATCT	6599
QY	661	ACAAAGAACCAAAACATGTTGCTCCGTCCTCAAAATACCTCCACCATCGCCACACCGCTCGG	720
DB	6600	ACGAAGNATCAACATGTTGCTCCGTCCTCAAAATACCTCCACCATCGCCACACCGCTCGG	6659
QY	721	GAGATCAAACTCAAGACACCCCACTGATGCCACAACTCAATACACGAGACCAAGC	780
DB	6660	GAGATCAAACTCAAGACACCCCACTGATGCCACAACTCAATACACGAGACCAAGC	6719
QY	781	AGTGATGATGAGGACCTCGCAACATCCGCTCAGGGTCCGGAGACGAGAACCCACACA	840
DB	6720	AATGATGATGAGGACCTCATACATCCGCTCAGGGTCCGGAGACGAGAACCCATAC	6779

Qy	841	ACTTCTGATGGCGTACCAAGCAAGGGCTTTTCATCAACAATGCGACCCACTCCCTCACCA	900
Db	6780	ACTTTCAGATGGCGTCACTTAAGCAAGGGCTTTTCATCAACAATGCGACCCACTCCCTCACCA	6839
Qy	901	CAACCAAGCAGCGCCACAGCAGGAGGAGAAACAACAACCAATTCCTCCCAAGATGCTGTGACT	960
Db	6840	CAACCAAGCAGCGCCACAGCAGGAGGAGAAACAACAACCAATTCCTCCCAAGGACTGTGACT	6899
Qy	961	GAACCTAGACAATAAACAACAATGCGCAACAACCGTCCATGCCCCCTCATACACTTACCACA	1020
Db	6900	GAACCAACAATAAACAACAATGCGCAACAACCGTCCATGCCCCCTCATACACTTACCACA	6959
Qy	1021	ATCTCTACTAACAACACTCCAAACAACAATTTCAGCACTCTCTCTGACCACTTACAAAAAC	1080
Db	6960	ATCTCTACTAACAACACTCCAAACAACAATTTCAGCACTCTCTCTGATTCACACAAAAAC	7019
Qy	1081	ACCACCAATGACAACAACAGAGACACAATCACTGAAATGAGCAAAACCACTGCCCCCTCG	1140
Db	7020	ACCACCAATGAGACAACAACAGAGACAGCACTGAAATGAGCAAAACCACTGCCCCCTCG	7079
Qy	1141	ATAACAACCCCTGCCCTCCAAAGGGGAAATCCCAACCAACAGCAAGAGCAGCAGCAAAAAA	1200
Db	7080	AAAACAACCCCTGCCCTCCAAAGGGGAAATCCCAACCAACAGCAAGAGCAGCAGCAAAAAA	7139
Qy	1201	GGCCCCGGCACAACGGCAGCAACCAACAGCAAAATGAGCAATTCACAGCTCTCCCCCACC	1260
Db	7140	GGCCCCCAGCAACGGCAGCAACCAATATGCAAAATGGCAATTAACCACTGCTCCCCCACC	7199
Qy	1261	CCAGCTCGACTGCAACACATCTTGTATATTTTCAGAGAAGCGA	1305
Db	7200	CCCAACCGGACCAACACATCTTGTATATTTTCAGAGAAGCGA	7244

RESULT 11

AF005733

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

AF005733

Marburg virus strain M/S.Africa/Johannesburg/1975/Ozolin glycoprotein precursor (GP) gene, complete cds.

AF005733

AF005733.1

GI:2459875

Lake Victoria marburgvirus
Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.
1 (bases 1 to 2046)

Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and Feldmann,H.

Variation in the glycoprotein and VP35 genes of Marburg virus strains

Virology 240 (1), 138-146 (1998)

98110148

9448698

22 (bases 1 to 2046)

Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.

Submitted (28-MAY-1997) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Building 15, Room SB611, Mail Stop G14, Atlanta, GA 30333, USA

Location/Qualifiers

1. .2046

/organism="Lake Victoria marburgvirus"

/mol_type="genomic RNA"

/strain="M/S.Africa/Johannesburg/1975/Ozolin"

/db_xref="taxon:11269"

1. .2046

/gene="Gp"

1. .2046

/gene="Gp"

/note="vriion surface glycoprotein"

/codon_start=1

/product="glycoprotein precursor"

/protein_id="AAC40458.1"

```

/db_xref="GI:2459876"
/translation="MRTTCLFTLSLILQIKTLPILEIASNNQPNVDSVCSGTLQKQ
EDVHLMGFTLSGQKVADEPLEASRWAFRTGVPKNVYETGEBAKTCYNIISVYDSFG
KSLLLDPTNVRVDPKCTIHHIQGNQPHAGIALHGWAFFLYDRIASITMYRGKVF
TEGNIIRAMTVNKTVMHVIISROQGYRHHMLTSTNKVWTSNSGNTQNDTCLGTLHEY
NSTENQTCAPSKTPPPPTAPRPIKPTSPDTATRLNTNPNNSDDDELTTSGSGSGEO
ELYTSDAVTQKLSSTMPPTSPSPQCTFOOQGNNTNHSQDATDELDTNTNTTAQPPPT
SHNTTITNTNTSKHNLTSAPPTNPTNPTQSMVTENEKTSAPKATLPPPTENPTT
EKSTNNKTPPTLEPTNTNGHTSPSTNSHTQHLYIFRRKRSILWREGDMFFLDG
LINAPDIFDPVNTKTIIFDESSSGASTBEDQHSNII SLTSLYSHTSNTAYSGEN
ENDCDELRLTWSVOREDDLAAGLSWILPLFGPIEGLTAGLIIKNNLVCRLRLANOT
AKSLDELRTVTEETREFSLINHWIADPELLTRWGGTCKVLGPDCIIGIEDLSRNIISQI
DQIKDEQKRGTVGLGGRKWTWSDMGVNLGLILLLSLAVLSALSCICRIFTKYIG"
1..81
/sig_peptide
/gene="Gp"
mat_peptide
/length=81
82..2043
/gene="Gp"
/misc_feature
/product="glycoprotein"
/length=2043
1930..2019
/gene="Gp"
/note="encodes transmembrane anchor peptide"
ORIGIN
Query Match          59.7%; Score 1098.6; DB 14; Length 2046;
Best Local Similarity 50.1%; Pred. No. 0;
Matches 117; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 1 ATGAAGACCACTGTTTTCCTTATCAGTCCTTATCTTAATTCACGGGACCAAAAAATCTCCCC 60
DB 1 ATGAGGACTACATGCCCTCTTTATCAGTCCTTATCTTAATTCACGGGATAAAACTCTCCCT 60

QY 61 ATTTAGAGATAGTACTAGTAAATAATCAACCCCAAAAAATGTGGATTCGTCTCGGAAC 120
DB 61 ATTTTGAGATAGCCAGTAACGATCAACCCCAAAAAATGTGGATTCGTCTCGGAAC 120

QY 121 CTCAGAGACAGAGAGCTCCATCTGTATGGGATTCACACTGATGGGCGAAAAAGTTGCT 180
DB 121 CTCAGAAAAACAGAAAGATGTCATCTGTATGGGATTTACACTGAGCGGCGAGAAAGTTGCT 180

QY 181 GATTCCCTTTTGAGGATCCAAACGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAT 240
DB 181 GATTCCCTTTTGAGGATCCAAACGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAT 240

QY 241 GTTGAGTACACAGAGGGGAGGAAGCCAAACATGCTCAATATAAGTGTAAACGATCCC 300
DB 241 GTTGAGTACACGGAAGGGAAGAACCAACATGCTCAATATAAGTGTAAACGATCCC 300

QY 301 TCTGGAAAAATCCTTGCTGTAGATCCTCTCAACAAATCCGTGACTATCGAAATGCAAA 360
DB 301 TCTGGAAAAATCCTTGCTGTAGATCCTCTCCCAACAAACGTCGTGACTATCCTAAATGCAAA 360

QY 361 ACTATCCATCATATTCAAGTGTCAAAACCCCTCATGCAAGGGATCGCCCTTCATTTATGG 420
DB 361 ACTATCCATCATATTCAAGTGTCAAAACCCCTCATGCGCAGGGATTCGCCCTTCATTTATGG 420

QY 421 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAACATGTACCGAGGCAAGTCTTC 480
DB 421 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAACATGTATCCAGGCAAGTCTTC 480

QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGAAGTGTCAAAAAATGATTTCTCG 540
DB 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGAAGTGTCAAAAAATGATTTCTCG 540

QY 541 CGGCAAGGACAGGGTACCGTCAATGAATCTGACTTCTACTAATAAATATGTGACAAAGT 600
DB 541 AGGCAAGGACAGGGTACCGTCAATGAATCTGACTTCTACTAATAAATATGTGACAAAGT 600

QY 601 AGTAACGGAAACGAAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 660
DB 601 AGCAACGGAAACGAAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 660

QY 661 ACAAGGAACCAACATGTGCTCGTCCGTCGTAATAATACCTCCACACTGCCACAGCCCGTCG 720

```

```
Db 661 ACGAAGAACCAACATGTGCTCGGTCTTAAACACACCCCCACACCCGCCCGTCCG 720
Qy 721 GAGATCAAACTCAACAGCACCCCAACTGATGCGCACCACCAACTCAATACACGACCAAGC 780
Db 721 GAAATCAAAACCAACAGCACCCCAACCGATGCGCAGTACGATCAACACCAACCAAC 780
Qy 781 AGTGATGATGAGGACCTCCAAACATCCGCGCTCAGGCTCGGAGACGAGAACCCACACA 840
Db 781 AGTGATGATGAGGATCTCAACACATCCGCGCTCGGCGTCTGGGGAACAGAACTCTATAGC 840
Qy 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTCATCAACATGCCACCCACTCCCTCACA 900
Db 841 ACTTCTGATGCGGTCACTAAGCAAGGGCTTTCATCAACAAATGCCACCCACTCCCTCACC 900
Qy 901 CAACCAAGCACGCCCAACAGAGGAGAAACAACAACCAATCTCCCAAGATGCTGTGACT 960
Db 901 CAACCAAGCACGCCCAACAGAGGAGAAACAACAACCAATCTCCCAAGAGCTTACACT 960
Qy 961 GAACTAGACAAAATAACCAACTGCGACACCGTCCATGCCGCCCTCATAAACAATACCA 1020
Db 961 GAACTTGATAACCAACCAACAACTGCGACAAACCGCCACGCGCTCCCAACAACCAACA 1020
Qy 1021 ATCTCTACTAAACAACCTCCAAACACAACTTCAGCACTCTCTCTGACCACTTACAAAC 1080
Db 1021 ATCTCTACCAACAACACCTCCAAACACAACTTCAGCACTCTCTGCGCACCAACCAAA 1080
Qy 1081 ACCACCAATGACACACAGACGACCAATCACTGAAATGACAAACCAAGTCCGCCCTCG 1140
Db 1081 ACCACCAATGCCCAACACAAAGACATGCTGCTGAAATGAGAAACCAAGTCCGCCCTCG 1140
Qy 1141 ATAACAACCTCGCTCCAAACGGGAAATCCCAACAGAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 AAAGCAACCTCGCTCCAAACAGAAATCTACCAAGAGAGAGAGAGAGAGAGAGAGAG 1200
Qy 1201 GGGCCGCGCAACAGGACCAACAGCAACAAATAGAGATTCACAGTCTCTCCGCCACCC 1260
Db 1201 AGCCCCCACCACATTTGGAACCAATAACAACAAATGAGACATTTCACTAGTCTCTCC 1260
Qy 1261 CCAGCTCCAGTCGACAACTCTGTATATTTTCAGAGAAAGCGA 1305
Db 1261 CCCAAGTCAGTACACACATCTTATATATTTTCAGAGAGAAAGCGA 1305

RESULT 12
LOCUS AY358025
DEFINITION Marburg virus strain M/S.Africa/Johannesburg/1975/Ozolin, complete
sequence.
ACCESSION AY358025
VERSION AY358025.2 GI:34304751
KEYWORDS
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
VIRUSES: serNA negative-strand viruses; Mononegavirales;
Filooviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 19151)
AUTHORS Bowen,M.D., Thurman,K., Minor,E., Ibrahim,M.S., Meyer,R.F.,
Malfatti,S.A., Do,L.H., Smith,K.L., McCready,P.M. and Chain,P.S.G.
Direct Submission
TITLE Submitted (01-AUG-2003) Viral Sequencing Group, Lawrence Livermore
NATIONAL National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA
COMMENT On Aug 27, 2003 this sequence version replaced gi:33868618.
FEATURES
source
1..19151
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/S.Africa/Johannesburg/1975/Ozolin"
/db_xref="taxon:11269"
/notes="cRNA sequence"
primer_bind 1..19
gene 123..2210
/genes="NP"
CDS 123..2210
```

```
/gene="NP"
/notes="nucleoprotein"
/codon_start=1
/product="NP protein"
/protein_id="AAQ55255.1"
/db_xref="GI:33868619"
/translation="MDLHSLLELGTGTAPAHVRNKKVILFDTNHQVISCNQIIDAINS
GIDLDLLEGLLTLCVEHYNSDKDFNTSPIAKYLRDAGYEFEDVIKNAATRFVDV
IPNEPHYSPLIALKLTLESTESQGRIGLFLSCFLPLKLVGDRASIEKALROVTV
HOEGIVTVYNNHLLTGHMKVIFGLIRSSFILKFLVLIHQGVNLVTGHDAYSIIINSV
GOTRPSGLIVKTVLETLQXDSQVTLHPLVRTSKVNEVASFKQALSNLARHGEYA
PFARVLNSGINNLEHGLYFQLSALGAVATAHGSTLAGVNVGEQYQQLREAAHAEV
KLQRHHEQEIQAIAEDDEERKILEQFHLQKTEITHSQTALVLSKRSKRLAKLAABE
NIIVDDQGFQKQSNRVSQSFNDPTPEVTVOARPVNRPALTALPPVDDKIEHSTEDS
SSSSFDVNDPFDALLNEDEDTLDSVMPTSTTREFQGIPESPQSGDLDNSQSQKE
NSNTPIKKQFLRYQELPPVOEDSEYTTDSQSIIDQSGDNEQGVLDLPPPLYTQE
KROPIQHPAASSODPFGSIGDVNGDILLEPIRSPSSPSAPOEDTTRARAYELSPDPTN
YEDQXQMPQRVVTKKGTFLYPNDLLQTSPPESLVTALVEEYQNPVSAKELQADWPD
MSFDRHRHVAANL"
2964..3953
/gene="VP35"
2964..3953
/notes="putative minor component of RNA-dependent RNA-
polymerase"
/codon_start=1
/product="VP35 protein"
/protein_id="AAQ55256.1"
/db_xref="GI:33868620"
/translation="MWDSSYMQQVSEGLMTGKVPIDQVFGANPLEKLYKRRPKGTG
LQCFPLMSKATSTDDIWDQLIVKLTADLLIPINQISDIQSTLSEVTTKRVHIEIR
OQHEITVPKLTLEATISKGMEMLAKYDLHVTSTGRTTAPAAAFDAYLNEHGVPPP
OPATPKDLGVAQAQCKGTVMYKNATDAADKMSKVLSESETFSPKPNLSAKLALLF
THLPNDNTPPHILAQVLSKIAYKSGKCAFLDAPHQIILSEGENAQAALTSLSTFADP
LGVVPPVIRKVPQTVPRPCOKSLRAVPNPNTIDKGVYCVYSSEQGETRALIK"
4587..5498
/gene="VP40"
4587..5498
/notes="putative matrix protein"
/codon_start=1
/product="VP40 protein"
/protein_id="AAQ55257.1"
/db_xref="GI:33868621"
/translation="MASSNNTYMYLNPYPYADHGANOILIPADQLSNHQHGTTPNV
GDNLDDQFKGNVCHAFLEAIIDISAVNERTKGVPAWLPLGINSPEYPLAHTVAA
LLTSGYITQTTHNGQKQFVRNRLGTGIPAHPLMLREGNOAFIQNMVLPKNSTNQF
TYNLTNVLVSQKLPDDAWRPSKDKLIGNTMHPAVSIHPNLPPIVLPTVKKQAYRQHK
APENFPLNGFNRRQVVLAYANPTLSAI"
5960..8005
/gene="GP"
5960..8005
/notes="envelope transmembrane glycoprotein"
/codon_start=1
/product="GP protein"
/protein_id="AAQ55258.1"
/db_xref="GI:33868622"
/translation="MRTTCLFSLILIGIKITLPILEIASNDQPNQVDSVCSGTLOKT
EDVHLMGFTLSGQKVDASLEASKRWAFRTGVPKNVYTEGEAAKTCYNISVTPSG
KSLIDPPTNVDRDYPCKCTIHHIQGNPHAQIALHMGAFPLDRIASTTMYRGKVF
TEGNTAAMVKNKTHKMFISRQGGYRHMLTSTNKYTSNGTNDTDCGCTGLHEY
NSTKNOTCAPSKTTPPTTARPEIKPTSTPTDATLNTNPNNSDDEDLTTSGSGCEQ
ELYTTSDAVTQKGLSSTWPTTSPQPGTQGGNNTHSQDATDELDTNTTNTTAQPTT
SHNTTTSNTNTHSLNLSAPPONTNPNQTSWATENEKTSAPPKATLPTPTPTT
EKSTNNTKPTTMEPTNTGHTSPSTPNSTQHLIYFRKRSKLSLRMGDMFPFLDG
LINAPIDPVPNTKTIIDESSSSGASTEDQHASNSITLSLYKRPSTAGSGEN
ENDCDALRIVSQEDDLAAGLWIPFPGIEGLYTAGLIKQNNVLCRLRLANQI
AKSLELLRVTEERTFSLINRHAIDFLLTRWGTCKVLGPCCIGIELSNISBQI
DQIKDEQKEGTGWLGGKWTSDMGVLTNLGILLILLSTAVLIASCICRIFTKYIG"
8888..9733
/gene="VP30"
```



```
QY 961 GAACTAGACAAAATAACACAACCTGCAAAACCGTCCATGCGCCCTCATTAACACTACCAACA 1020
D 6920 GAACTTTGATAACCAACCAACAACTGCAAAACCGGCCCAACCGCCCTCCGCAACACACCAACCA 6979
QY 1021 ATCTCTACTAACACACACCTCCAAACACAACTTTCAGCACTCTCTCTGCGACCATTAACAAC 1080
D 6980 ATCTCCACCAACAAACACCTCCAAACACAACTTTCAGCACTCTCTCTGCGACCATTAACAAC 7039
QY 1081 ACCACCAATGACAAACACACAGAGCACAATCACTGAAATAATGAGCAAAACCAAGTCCCGCTCG 1140
D 7040 ACCACCAATCCCAACACACAAAGCATGGCCACTGAAATGAGAAACCAAGTCCCGCTCG 7099
QY 1141 ATAACAACCTCGCTCCCAACCGGAATCCCAACACAGCAAGAGAGACCAAGAGCAAGCAAAA 1200
D 7100 AAAGCAACCTCGCTCCCAACAGAAATCTACCAACAGAAAGAGACCAACCAATACAAA 7159
QY 1201 GCGCCGCGCACCAACGCGCACCAACAGCAAAATGAGCAAAACCAAGTCCCGCTCG 1260
D 7160 AGCCCCACCAATGGAACCAATACAAATGAGCAATTTCACTAGTCCCTCCAC 7219
QY 1261 CCAGCTCGACTGCAACCAATCTTGATATATTTTCAAGAAAGCGA 1305
D 7220 CCNACTCGACTACACCAATCTTATATTTTCAGGAGGAACGA 7264

RESULT 13
AX717719
LOCUS AX717719 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 7 from Patent WO02079239.
ACCESSION AX717719
VERSION AX717719.1 GI:29890728
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Grogan,C.C.; Hevey,M.C. and Schmaljohn,A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 7 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source
1..2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/noE="chimeric molecule between Marburg virus strain
Raven Glycoprotein 1 and Marburg virus strain Musoke
Glycoprotein 2"

ORIGIN
Query Match 42.4%; Score 780; DB 6; Length 2046;
Best Local Similarity 74.8%; Pred. No. 6.7e-225;
Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;
QY 1 ATGAAGACACATGTTTCCTTATCACTGCTTATTAATTCAGAGGACAAAAATCTCC 60
D 1 ATGAAGACATATATTTCTGATTAGTCTCAITTTAATCCAAAGTATAAAACTCTCCT 60
QY 61 ATTTTAGATAGTAGTATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGA 120
D 61 GTTTTAGAAATTTGCTAGTAAACGCCAACCTCAAGATGTAGATTCAAGTGTGCTCCGGA 120
QY 121 CTCGAGACAGAACAGCTGCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
D 121 CTCGAGACAGAACAGATGTTTCACTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCTCTTTGGAGGCATCCAGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
D 181 GATTCCTCTTTGGAGGCATCTAAACGATGGCTTTTCAGGACAGGTGTCTCCCAAGAAC 240
QY 241 GTTGTAGTACAGAGGGGGAGGAGGCAACATGCTACAAATATAAGTGTACAGGATCCC 300
D 241 GTTGTAGTATACGGAAGGAAGGCAACATGTTTACAAATATAAGTGTACAGACCTC 300
```

RESULT 14

AX717725

LOCUS

AX717725

2046 bp

DNA

linear

PAT 15-APR-2003

DEFINITION	Sequence 13 from Patent WO02079239.	
ACCESSION	AX717725	
VERSION	AX717725.1	GI:29890731
KEYWORDS	Lake Victoria marburgvirus	
SOURCE	Lake Victoria marburgvirus	
ORGANISM	Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.	
REFERENCE	1	
AUTHORS	Grogan,C.C., Hevey,M.C. and Schmaljohn,A.L.	
TITLE	Chimeric filovirus glycoprotein	
JOURNAL	Patent: WO 02079239-A 13 10-OCT-2002;	
U.S. Army Medical Research Institute of Infectious Diseases (US)		
FEATURES	Location/Qualifiers	
source	1..2046	
	/organism="Lake Victoria marburgvirus"	
	/mol_type="unassigned DNA"	
	/strain="Raven"	
	/db_xref="taxon:11269"	
	/note="chimeric molecule between Marburg virus Glycoprotein 1 and Marburg virus Glycoprotein 2"	
ORIGIN		
Query Match	42.4%;	Score 780; DB 6; Length 2046;
Best Local Similarity	74.8%;	Pred. No. 6.7e-225;
Matches	978;	Conservative 0; Mismatches 330; Indels 0; Gaps 0;
QY	1	ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAAATTCAGGGGACAAAAATCTCCCC 60
Db	1	ATGAAGACCACATATATTTCTGATAGTCTCATTTTAAATTCNAAGTATAAAACTCTCCCT 60
QY	61	ATTATTAGATAGCTAGTAAATTAATCAACCCCAAAATGGGATTCGGTATGCTCCGGAAT 120
Db	61	GTTTTAGAAATTCGTAGTAACGCAACCTCAAGATGTAGATTCAGTGTGCTCCGGAACC 120
QY	121	CTCCAGACAGAGAGCTCCATCTGTGGGATTCACACTGAGTGGGCAAAAGTTCCT 180
Db	121	CTCCAAAGACAGAGAGATGTTCTATCTGATGGGATTTACACTGAGTGGGCAAAAGTTCCT 180
QY	181	GATTCCTCTTGGAGGATCCAGGAGTGGGCTTTCAGGACAGGTGTACTCCCAAGAT 240
Db	181	GATTCCTCTTGGAGGATCTAAGCAGTGGGCTTTCAGGACAGGTGTCTCCCAAGAAC 240
QY	241	GTGTAGTACAGAGGGGAGGAGCAAAAATGCTACAAATATAGTGTAAACGGATCCC 300
Db	241	GTGTAGTATACGGAAGGAGAGCAAAAATGTTTACAAATATAGTGTAAACGAGCCT 300
QY	301	TCTGGAATCTCTGCTGTAGATCTCTACCAACATCCGTCGACTATCCGGAATGCAAA 360
Db	301	TCTGGAATATCTCTGCTGTAGATCTCTCCAGTAATATCCGGATTTACCCCTAAATGTAAA 360
QY	361	ACTATCCATCATATTCAAGGTCAAAACCTCATGCACAGGGATGCCCTTCATTTATGG 420
Db	361	ACTGTTTCATCATATTCAAGGTCAAAACCTCATGCACAGGGATGGCCCTTCATTTGGG 420
QY	421	GGAGCATTTTCTGTATGATCGATGCTCCACAAATATGACGAGGCAAAAGTCTTC 480
Db	421	CGGGCATTTTCTGTATGATCGGTGCTCTACAAATATGACGAGGCAAAAGTCTTC 480
QY	481	ACTGAAGGAAACATAGCAGCTATGATTTGCAATAGACAGTGCACAAATGATTTTCG 540
Db	481	ACTGAAGGAAATATAGCAGCTATGATTTGTAATAGACAGTTCACAGAAATGATTTTCT 540
QY	541	CGGCAAGGACAGGGTACCGTCAATGAAATCTGACTTCTACTAATAATATTTGGCAAGT 600
Db	541	AGGCAAGGACAGGGTATCGTCAATGAAATCTGACTTCCACCAATATAATTTGGCAAGC 600
QY	601	AGTAACGGAACGCAAAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAAATCT 660
Db	601	AGCAATGAAACGCAAGAAATGATACGGGATGTTTGGCATCTCTCAAGAAATACAACTCC 660
QY	661	ACAAAGAACAAACATGTGCTGGTCCAAAATACCTTCACCACTGCCACAGCCCGTCCG 720
Db	661	ACAAACAAATCAAAATGCGCTCCATCTCTTAAACCTTCATCCCTGCCACAGTAACCTCG 720
QY	721	GAGATCAAACTCAACAGCAGCCCACTGATGCCCAAACTCAATACACGAGCAACCAAGC 780
Db	721	AGCATTCACCTCAAAATCTCAAAATTAATCTGTAATCTGGAATCTATGAACCAAGT 780
QY	781	AGTATGATGAGGACCTCGCAACATCCGGCTCAGGTCGGGACGAGACGAGAACCCACACA 840
Db	781	AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
QY	841	ACTTCTGATGCGTCCACCAAGCAGGGCTTTTCATCAACAAATGCCACCACTCCCTCCACCA 900
Db	841	ACTCTTAATGATGACCTGACAGAGAAACAATGCTCAACAATATTTGTCACCTCTTCACTA 900
QY	901	CAACCAAGCAGCCACAGCAGAGGAGAAACAACAACCAATATTCACAAAGATGCTGTGACT 960
Db	901	CATCCAGCAGCTCAACAATGAGCAAAACAGTACGAATCTTCCGACATGCTGTAAT 960
QY	961	GAACTAGACAAATTAACAACTGCAACAGCGTCCATGCCCTCTCAATAAAGTCTTAATACA 1020
Db	961	GAGCAAAATGGAACCGACCCCAACAACCAACAGCAGCAACGCTCTCAACAATTAATACA 1020
QY	1021	ATCTCTACTAAACACACCTCCAAAACACAACTTCAGCACTCTCTGTGCACCACTTACAAAAC 1080
Db	1021	ACTCCACCTATACACTCTCAAGTACAACTCAGTACTCTTCCCTCCACCGCAAC 1080
QY	1081	ACCACCAATGACACACAGCAGCACAATCACTGAAATAGCAAAACAGTGCCTCCCTCG 1140
Db	1081	ATCACAATATGATACACAACGTGAACTAGCAGAAAGCGAACAACCAATGCTCAGTTG 1140
QY	1141	ATAACAACCTGCTCCGAAACGGGAAATCCCAACACAGCAAGAGCAGCAGCAAGAAAAA 1200
Db	1141	AACAACCTCTAGATTCACAGAAATCCCAACAGCAAGAACAGCAAGCAAGCAACCAACC 1200
QY	1201	GCCCCCGCACAAAGCGGACCAACACAGCAAAATGAGCAATTTTACCAGTCTCTCCCCCACC 1260
Db	1201	AACATCATCATGACAGATCAGATATACAGCAAAACACCCCAACCAATTTCTTCGGAT 1260
QY	1261	CCCAGCTCGACTGCACACATCTTTGTATATTTTCAGAGAAAGCGATCG 1308
Db	1261	TCTAGTCGACAAACCGGCTCTCTATATATTTAGAAAGAAAGATCG 1308
RESULT 15		
AF005734		
LOCUS	AF005734	2046 bp RNA linear VRL 23-JAN-1998
DEFINITION	Marburg virus strain M/Kenya/Kitum Cave/1987/Ravn glycoprotein precursor (GP) gene, complete cds.	
ACCESSION	AF005734	
VERSION	AF005734.1	GI:2459877
KEYWORDS	Lake Victoria marburgvirus	
SOURCE	Lake Victoria marburgvirus	
ORGANISM	Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.	
REFERENCE	1 (bases 1 to 2046)	
AUTHORS	Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and Feldmann,H.	
TITLE	Variation in the glycoprotein and VP35 genes of Marburg virus strains	
JOURNAL	Virology 240 (1), 138-146 (1998)	
MEDLINE	98110148	
PUBMED	9448698	
REFERENCE	2 (bases 1 to 2046)	
AUTHORS	Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.	
TITLE	Direct Submission	
JOURNAL	Submitted (29-MAY-1997) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Building 15, Room SB611, Mail Stop G14, Atlanta, GA 30333, USA	
FEATURES	Location/Qualifiers	
source	1..2046	
	/organism="Lake Victoria marburgvirus"	
	/mol_type="genomic RNA"	

```
/strain="M/Kenya/Kitum Cave/1987/Ravn"
/db_xref="taxon:11269"
1. .2046
/gene="GP"
1. .2046
/gene="GP"
/notes="virion surface glycoprotein"
/codon_start=1
/product="glycoprotein precursor"
/protein_id="AAC40459.1"
/db_xref="GI:2459878"
/translation="MKTLYPLSLILQSIKTLPLVLEIASNSQPDVDSVCSGLQKT
EDVHMLGSLQKQVADSPLEASKRWAFRTVPKNVEYTEGEERKTCINISVDPFG
KSLLDPPSNIRDYPKCTVHHIOGONPHOAGIALHWGAFFLYDRVASTMYRQVF
TEGNTAAIVNKVHRMIFSRQOQYRHMNLTSTNKWYTSNETDRNDTGFGLQY
NSTNQTCPPSLPSTVTPSIHSTNTQAKSGTWNPSDDDEDLIMISGSGSEQ
GPHYTLNVVTEQKQSSSTLSTPSLHSTSOHEONSTNP SRHATVETHNCTDPTOPATL
LNNVTYTYNTLKNLSTPSPPTRNITNNDQRELASEQTNAGLNTLLDPTENPTT
GQDTSNTIIMTSDITSKPTNSPSSPTTPPIYFRKGRSIFWKEGDIFFPDG
LINTEDFDPIINTETIDFESFNTSNEOHTPPNISLTSFSPDKNXGDTAYSGEN
ENDCDALRINSVQDDLAAGLSWTFPGPGIEGLYTAGLKNQNLVCLRLRLANQT
AKSLLELLRVVTEERTFSLINRHAIDFLLTRWGTCKVLGPDCCIGIEDLSKNTSEQI
DKIRDEQKEETGWLGGKWTSDWGLTNLGLILLLSIAVLIALSCIRFTKYIG"
1. .81
/sig_peptide
/gene="GP"
82. .2043
/mat_peptide
/gene="GP"
/product="glycoprotein"
1930. .2019
/misc_feature
/gene="GP"
/notes="encodes transmembrane anchor peptide"

ORIGIN
Query Match
Best Local Similarity 42.2%; Score 777; DB 14; Length 2046;
Matches 975; Conservative 0; Mismatches 330; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATGTTTCTTATAGTCTTATCTTAAATTCAGGGACCAAAAATCTCCCC 60
DB 1 ATGAAGACCATATATTTCTGATTAGTCTCAATTTTAAATCCAAAGTATAAAAATCTCCCT 60
QY 61 ATTTAGAGATAGTAGTAATAATCAACCCCAAAATGGGATTCGGTATGCTCCCGAAT 120
DB 61 GTTTTGAAGAAATGCTAGTAACGCCAACCTCAAGATGTAGATTCAAGTGTGCTCCGGNAAC 120
QY 121 CTCGAGACAGAGACGCTCATCTGATGGATTACACCTGAGTGGGCAAAAAGTTGCT 180
DB 121 CTCGAAAAGACAGAGATGTTTCTGATGGATTACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTGGAGCATCCAAGCGATGGGCTTTTCAAGACAGGTGTACCTCCCAAGAA 240
DB 181 GATTCCCTTTGGAGCATCTAAACGATGGGCTTTTCAAGACAGGTGTCTCTCCCAAGAAC 240
QY 241 GTTGAGTACAGAGGGGAGAGAGCCAAAACATGCTACAAATATAAGTGTACAGGATCCC 300
DB 241 GTTGAGTATACGGAAGGAGAGAGCCAAAACATGTTTACAAATATAAGTGTAAACAGACCT 300
QY 301 TCTGGAATATCTGCTGTTAGATCTCTCTACCAATCCGTGACTATCCGAAATGCAAA 360
DB 301 TCTGGAATATCTGCTGTTAGATCTCTCTACCAATCCGTGACTATCCGAAATGTA 360
QY 361 ACTATCCATCATATTCAGGTCAAAACCTCATGCAAGGGATCGCCCTTCATTTATGG 420
DB 361 ACTGTTTCATCATATTCAGGTCAAAACCTCATGCAAGGGATTCGCCCTCCATTTGG 420
QY 421 GGAGCAATTTTCTGTATGATCGCATTCGCTCCACAAAGTGTACCGAGGCAAGTCTTC 480
DB 421 GGGGCAATTTTCTGTATGATCGCTTCCTACCAAGTGTACCGAGGCAAGTCTTC 480
QY 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTTCTCG 540
DB 481 ACTGAAGGAAATATAGCAGCTATGATTGTATATAGACAGTTCACAGATGATTTTCT 540
```

Search completed: August 1, 2005, 02:43:41
Job time : 7917.14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:55:23 ; Search time 964.473 Seconds
(without alignments)
11299.693 Million cell updates/sec

Title: US-10-066-506A-3

Perfect score: 1841

Sequence: 1 atgaagaccacatgttctct.....tttgttttttagtgaattc 1841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: _geneseqn1980s:*
- 2: _geneseqn1990s:*
- 3: _geneseqn2000s:*
- 4: _geneseqn2001as:*
- 5: _geneseqn2001bs:*
- 6: _geneseqn2002as:*
- 7: _geneseqn2002bs:*
- 8: _geneseqn2003as:*
- 9: _geneseqn2003bs:*
- 10: _geneseqn2003cs:*
- 11: _geneseqn2003ds:*
- 12: _geneseqn2004as:*
- 13: _geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1841	100.0	1841	10	ABT13451 DNA encod
2	1308	71.0	2046	10	ABT13452 DNA encod
3	1303.4	70.8	2103	8	ACC57764 Marburg v
4	1303.4	70.8	2164	4	AAD04042 Marburg v
5	1303.4	70.8	7005	10	ACC71551 VRC6702 (
6	1303.4	70.8	7778	10	ACC71550 VRC6701 (
7	1303.4	70.8	8256	10	ACC71552 VRC6710 (
8	1303.4	70.8	11460	3	AAX87211 VBS repli
9	1287.4	69.9	2100	3	AAX51039 Marburg v
10	1274	69.2	2051	10	ABT13454 DNA encod
11	1148.2	62.4	19112	13	ADM48345 Marburg v
12	780	42.4	2046	10	ABT13453 DNA encod
13	780	42.4	2046	10	ABT13456 DNA encod
14	773.8	42.0	2247	4	AAD04043 Marburg v
15	688	37.4	6902	10	ACC71562 VRC6703 (p
16	540	29.3	2039	10	ABT13455 DNA encod
17	527.6	28.7	2224	3	AAX51038 Ebola vir
18	527.6	28.7	7272	2	AAX59393 Plasmid p
19	527.6	28.7	7285	2	AAX89797 DNA of pV
20	527.6	28.7	9106	12	ADF85419 Ebola vir

21	527.6	28.7	18959	13	ADM48337	Adm48337 Zaire ebo
22	527.6	28.7	18959	13	ADM48364	Adm48364 Zaire ebo
23	527.6	28.7	18959	13	ADM48355	Adm48355 Zaire ebo
24	526	28.6	2298	3	AAX87189	Aax87189 Ebola vir
25	526	28.6	2298	4	RAF76953	RAF76953 Ebola vir
26	526	28.6	2298	12	ADL27460	Adl27460 Nucleotid
27	526	28.6	6467	10	ACC71530	Acc71530 VRC6052 (
28	526	28.6	6624	10	ACC71523	Acc71523 VRC6002 (
29	526	28.6	7154	10	ACC71521	Acc71521 VRC6000 (
30	526	28.6	7188	10	ACC71522	Acc71522 VRC6001 (
31	526	28.6	7272	2	AAX89798	Aax89798 DNA of pV
32	526	28.6	7285	2	AAX59392	Aax59392 Plasmid p
33	526	28.6	8439	10	ACC71548	Acc71548 VRC6603 (
34	525.6	28.5	2172	4	AAD04041	Aad04041 Ebola vir
35	525.6	28.5	6561	10	ACC71524	Acc71524 VRC6003 (
36	451.8	24.5	6914	10	ACC71529	Acc71529 VRC6008 (
37	451.8	24.5	8199	10	ACC71549	Acc71549 VRC6604 (
38	451.8	24.5	10783	10	ACC71545	Acc71545 VRC6600 (
39	420	22.8	7106	10	ACC71528	Acc71528 VRC6007 (
40	336.2	18.3	7044	10	ACC71527	Acc71527 VRC6006 (
41	324.6	17.6	7002	10	ACC71536	Acc71536 VRC6300 (
42	324.6	17.6	7023	10	ACC71541	Acc71541 VRC6351 (
43	324.6	17.6	7036	10	ACC71537	Acc71537 VRC6301 (
44	324.2	17.6	7001	2	AAX89795	Aax89795 DNA of pV
45	324.2	17.6	7003	2	AAX59390	Aax59390 Plasmid p

ALIGNMENTS

RESULT 1	
ABT13451	
ID	ABT13451 standard; DNA; 1841 BP.
XX	
AC	ABT13451;
XX	
DT	30-JAN-2003 (first entry)
XX	
DE	DNA encoding a chimeric filovirus protein MBGV-GP1/EBOV-GP2.
XX	
KW	Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW	immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX	
OS	Unidentified.
OS	Chimeric.
XX	
PN	WO200279239-A2.
XX	
PD	10-OCT-2002.
XX	
PF	31-JAN-2002; 2002WO-US003339.
XX	
PR	31-JAN-2001; 2001US-0267522P.
XX	
PA	(USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX	
PI	Grogan CC, Hevey MC, Schmaljohn AL;
XX	
DR	WPI; 2003-040651/03.
XX	
DR	P-P5DB; ABJ18473.
XX	
PT	New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT	useful for inducing an immune response against infection of different
PT	filoviruses, specifically against both Ebola and Marburg viruses.
XX	
PS	Claim 14; Page 70-72; 9app; English.
XX	
CC	The invention relates to a chimeric filovirus glycoprotein (GP) protein
CC	comprising GP1 and GP2, where GP1 is from a filovirus different than that
CC	of GP2. The chimeric filovirus GP protein is useful for inducing an
CC	immune response against infection of different filoviruses, specifically
CC	against both Ebola and Marburg viruses by being used as a vaccine. This
CC	polynucleotide sequence represents a DNA encoding a chimeric filovirus

CC	protein of the invention									
XX	Sequence 1841 BP; 589 A; 500 C; 361 G; 391 T; 0 U; 0 Other;									
SQ	Query Match 100.0%; Score 1841; DB 10; Length 1841; Best Local Similarity 100.0%; Pred. No. 0; Matches 1841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	ATGAAGACCA	CATGTTTCCCTTATCAGTCTTATCTTAATTTCAAGGGACAAAAATCTCCCC	60						
Db	1	ATGAAGACCA	CATGTTTCCCTTATCAGTCTTATCTTAATTTCAAGGGACAAAAATCTCCCC	60						
Qy	61	ATTTTAGAGAT	AGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	120						
Db	61	ATTTTAGAGAT	AGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	120						
Qy	121	CTCCAGAGAC	GAGAGAGCTCCATCTGATGGGATTTACACTGAGTGGGCAAAAGTTGCT	180						
Db	121	CTCCAGAGAC	GAGAGAGCTCCATCTGATGGGATTTACACTGAGTGGGCAAAAGTTGCT	180						
Qy	181	GATTCCTCTT	GGAGGCATCCAAGGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAAT	240						
Db	181	GATTCCTCTT	GGAGGCATCCAAGGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAAT	240						
Qy	241	GTGTAGTACA	GAGGGGAGGAAGCAAAATGCTACAAATATAAGTGTAAAGGATCCC	300						
Db	241	GTGTAGTACA	GAGGGGAGGAAGCAAAATGCTACAAATATAAGTGTAAAGGATCCC	300						
Qy	301	TCGTGNAAT	CTCTGTTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCCAA	360						
Db	301	TCGTGNAAT	CTCTGTTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCCAA	360						
Qy	361	ACTATCCAT	CATATTAAGGTCAAAACCTCATGCACAGGGATCGCCCTTCATTTATGG	420						
Db	361	ACTATCCAT	CATATTAAGGTCAAAACCTCATGCACAGGGATCGCCCTTCATTTATGG	420						
Qy	421	GGAGCAAT	TTTTTCTGTATGATCGCATGCTCCCAACAATGTACCGAGGCAAAATCTTC	480						
Db	421	GGAGCAAT	TTTTTCTGTATGATCGCATGCTCCCAACAATGTACCGAGGCAAAATCTTC	480						
Qy	481	ACTGAAGGA	ACATAGCAGCTATGTTGTCATATAGACAGTGCACAAATGATTTCTCG	540						
Db	481	ACTGAAGGA	ACATAGCAGCTATGTTGTCATATAGACAGTGCACAAATGATTTCTCG	540						
Qy	541	CGCAAGGACA	AGGGTACCGTCTATATGAATCTGACTTTCTACTAATAATATTGGCAAGT	600						
Db	541	CGCAAGGACA	AGGGTACCGTCTATATGAATCTGACTTTCTACTAATAATATTGGCAAGT	600						
Qy	601	AGTAACGGAA	CGCAACGAAATGACACTGGATGTTTCGGGCTCTTCAAGAAATACAATCT	660						
Db	601	AGTAACGGAA	CGCAACGAAATGACACTGGATGTTTCGGGCTCTTCAAGAAATACAATCT	660						
Qy	661	ACAAAGAAC	CAATGCTCGTCCCAAAATACCTCCACCTGCGCCACAGCCGCTCG	720						
Db	661	ACAAAGAAC	CAATGCTCGTCCCAAAATACCTCCACCTGCGCCACAGCCGCTCG	720						
Qy	721	GAGATCAAA	CTCACAAGACCCCACTGATGCCACCAAACTCAATACCAAGCCCAAGC	780						
Db	721	GAGATCAAA	CTCACAAGACCCCACTGATGCCACCAAACTCAATACCAAGCCCAAGC	780						
Qy	781	AGTGATGAT	GAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA	840						
Db	781	AGTGATGAT	GAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA	840						
Qy	841	ACTTCTGAT	GCGGTCAACAGCAAGGGCTTTTATCAATGCGCACCACTCCCTCACC	900						
Db	841	ACTTCTGAT	GCGGTCAACAGCAAGGGCTTTTATCAATGCGCACCACTCCCTCACC	900						
Qy	901	CAACCAAGC	AGCCACAGAGGAGGAAACACACAAACCATTTCCCAAGATGCTGTGACT	960						
Db	901	CAACCAAGC	AGCCACAGAGGAGGAAACACACAAACCATTTCCCAAGATGCTGTGACT	960						
Qy	961	GAACTAGACA	AAAAATAACAACTGCAACACCGTCCATGCCCTCCATTAACACTACCA	1020						

Db	961	GAAC	TAGACAAAAATAACAACTGCAACACCGTCCATGCCCTCTATAACTACCACA	1020						
Qy	1021	ATCTCT	TACTTAACAACACCTCCAAAACACAACTTTCAGCACTCTCTCTGACCACTTACAAAAAC	1080						
Db	1021	ATCTCT	TACTTAACAACACCTCCAAAACACAACTTTCAGCACTCTCTCTGACCACTTACAAAAAC	1080						
Qy	1081	ACACCA	ATGACAAACACACAGAGCACAATCTGAAAAATGAGCAAAACGATGCCCCCTCG	1140						
Db	1081	ACACCA	ATGACAAACACACAGAGCACAATCTGAAAAATGAGCAAAACGATGCCCCCTCG	1140						
Qy	1141	ATPACA	ACCCCTCGCTCCCAACGGGAATCCCAACAGCAAAAGAGACACAGCAGCAAAAAA	1200						
Db	1141	ATPACA	ACCCCTCGCTCCCAACGGGAATCCCAACAGCAAAAGAGACACAGCAGCAAAAAA	1200						
Qy	1201	GGCCCC	CGCCACACCGCACCAACACGACAAATGAGCAATTTCCACAGTCCTCCCCCACC	1260						
Db	1201	GGCCCC	CGCCACACCGCACCAACACGACAAATGAGCAATTTCCACAGTCCTCCCCCACC	1260						
Qy	1261	CCAGCT	CGACTGCAACAACATCTTGTATATTTTTCAGAGAAAGCGATTCGGCAATTTGTCAAT	1320						
Db	1261	CCAGCT	CGACTGCAACAACATCTTGTATATTTTTCAGAGAAAGCGATTCGGCAATTTGTCAAT	1320						
Qy	1321	GCTCAAC	CCCAATGCAACCTTAATTTACTGACTACTCAGGATCAAGGTGCTGCA	1380						
Db	1321	GCTCAAC	CCCAATGCAACCTTAATTTACTGACTACTCAGGATCAAGGTGCTGCA	1380						
Qy	1381	ATCGGACT	GGCTGGATACATATTTCCGGCCACGACGAGGGAATTTTACATAGAGGG	1440						
Db	1381	ATCGGACT	GGCTGGATACATATTTCCGGCCACGACGAGGGAATTTTACATAGAGGG	1440						
Qy	1441	CTAATG	CACAATCAAGATGGTTTAACTCTGTGGTTTGAACAGCTGGCCAAACGAGACGACT	1500						
Db	1441	CTAATG	CACAATCAAGATGGTTTAACTCTGTGGTTTGAACAGCTGGCCAAACGAGACGACT	1500						
Qy	1501	CAAGCT	CTTCAACTGTTCTGAGAGCCCAACTGAGCTACGACACTTTTCAATCTCTCAAC	1560						
Db	1501	CAAGCT	CTTCAACTGTTCTGAGAGCCCAACTGAGCTACGACACTTTTCAATCTCTCAAC	1560						
Qy	1561	CGTAAG	GGCAATGATTTCTTCTGTCAGCGATGGGGCGGCACATGCCACATTTCTGGGACCG	1620						
Db	1561	CGTAAG	GGCAATGATTTCTTCTGTCAGCGATGGGGCGGCACATGCCACATTTCTGGGACCG	1620						
Qy	1621	GACTGCT	GATCGAACCATGATTTGGACCAAGAACATAAACAGCAAAATTCATCAGATT	1680						
Db	1621	GACTGCT	GATCGAACCATGATTTGGACCAAGAACATAAACAGCAAAATTCATCAGATT	1680						
Qy	1681	ATTCAT	GATTTTGTGATAAAAACCTTCCGGACCGAGGGGCAATGACAAATTTGGTGACA	1740						
Db	1681	ATTCAT	GATTTTGTGATAAAAACCTTCCGGACCGAGGGGCAATGACAAATTTGGTGACA	1740						
Qy	1741	GGATG	GAGACAAATGGATACCGCAGGTATTGGAGTTTACAGGGGTTATAATTTGCAGTTATC	1800						
Db	1741	GGATG	GAGACAAATGGATACCGCAGGTATTGGAGTTTACAGGGGTTATAATTTGCAGTTATC	1800						
Qy	1801	GCTTTAT	CTGTATATGCAAAATTTGTCTTTTAGTTGAATTC	1841						
Db	1801	GCTTTAT	CTGTATATGCAAAATTTGTCTTTTAGTTGAATTC	1841						

RESULT 2

ABT13452

ID

ABT13452 standard; DNA; 2046 BP.

XX

ABT13452;

AC

30-JAN-2003 (first entry)

XX

DNA encoding a chimeric filovirus protein MUS-GP1/RVN-GP2.

XX

Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;

XX

Immune response; Ebola; Marburg virus; vaccine; gene; ds.

XX

RESULT 2
ABT13452
ID ABT13452 standard; DNA; 2046 BP.
XX
AC ABT13452;
XX
DT 30-JAN-2003 (first entry)
XX
DE DNA encoding a chimeric filovirus protein MUS-GP1/RVN-GP2.
XX
KW Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
XX immune response; Ebola; Marburg virus; vaccine; gene; ds.

DB 961 GAACTAGACAAAAATAACAACTGCACAAACCGTCCATGCCCCCTCATAACTACCA 1020
QY 1021 ATCTCTACTTAAACAACACCTCCAAAACAACTTTCAGCACTCTCTCTGCACTTACAAAAC 1080
DB 1021 ATCTCTACTTAAACAACACCTCCAAAACAACTTTCAGCACTCTCTCTGCACTTACAAAAC 1080
QY 1081 ACCACCAATGACAAACACACAGAGCACAATCACTGAAATAGCAAAACAGTGCCCTCTG 1140
DB 1081 ACCACCAATGACAAACACACAGAGCACAATCACTGAAATAGCAAAACAGTGCCCTCTG 1140
QY 1141 ATAAACACCTGCTCCAAACGGGAATCCCAACAGCAAGAGAGCAGCAGCAAAAAA 1200
DB 1141 ATAAACACCTGCTCCAAACGGGAATCCCAACAGCAAGAGAGCAGCAGCAAAAAA 1200
QY 1201 GGCCCGCCACAAACGGCAACCAACACGCAAAATGAGCAATTTTCCAGTCTCTCCCCCACC 1260
DB 1201 GGCCCGCCACAAACGGCAACCAACAGCAAAATGAGCAATTTTCCAGTCTCTCCCCCACC 1260
QY 1261 CCCAGCTCGACTGCACAACTCTTGTATATTTTCAAGAAAGCGATCGGCAATTTGTCAT 1320
DB 1261 CCCAGCTCGACTGCACAACTCTTGTATATTTTCAAGAAAGCGATCGGCAATTTGTCAT 1320
QY 1321 GCTCAACCCAAATGCAACCTTAATTTTACATTTACTGGACTCTCAGGATGAGAGTCTCA 1380
DB 1321 GCTCAACCCAAATGCAACCTTAATTTTACATTTACTGGACTCTCAGGATGAGAGTCTCA 1380
QY 1381 ATCGGACTGGCTGATACCATATTTTTCGGGCGAGCAGCGAGGGAAATTTACATAGAGGG 1440
DB 1381 ATCGGACTGGCTGATACCATATTTTTCGGGCGAGCAGCGAGGGAAATTTACATAGAGGG 1440
QY 1441 CTAATGCACAATCAAGATGTTTAAATCTGTGGTTGAGACAGCTGGCCAAACGAGACGACT 1500
DB 1441 CTAATGCACAATCAAGATGTTTAAATCTGTGGTTGAGACAGCTGGCCAAACGAGACGACT 1500
QY 1501 CAAGCTCTTCAACTGTTCTCGAGAGCCAACTGAGCTACGCACTTTTCAATCTCTCAAC 1560
DB 1501 CAAGCTCTTCAACTGTTCTCGAGAGCCAACTGAGCTACGCACTTTTCAATCTCTCAAC 1560
QY 1561 CGTAAGGCAATTTGATTTCTGCTGAGGATGGGCGGCACATGCCACATTTCTGGGACCG 1620
DB 1561 CGTAAGGCAATTTGATTTCTGCTGAGGATGGGCGGCACATGCCACATTTCTGGGACCG 1620
QY 1621 GACTGCTGATCGAAACCACTGATTTGGACCAAGAACATAACAGACAAATTTGATCAGATT 1680
DB 1621 GACTGCTGATCGAAACCACTGATTTGGACCAAGAACATAACAGACAAATTTGATCAGATT 1680
QY 1681 ATTCAATGATTTTGTGATATAAAACCTTCCGACACGAGGGGCAATGCAATTTGGTGACA 1740
DB 1681 ATTCAATGATTTTGTGATATAAAACCTTCCGACACGAGGGGCAATGCAATTTGGTGACA 1740
QY 1741 GGATGGAGCAATGATACCGGCAAGTATTTGAGTTTACAGGGCTTATAATTTGCACTTATC 1800
DB 1741 GGATGGAGCAATGATACCGGCAAGTATTTGAGTTTACAGGGCTTATAATTTGCACTTATC 1800
QY 1801 GCTTTATTTCTGTATATGCAAAATTTGTCTTTTGTAGTTGAATTC 1841
DB 1801 GCTTTATTTCTGTATATGCAAAATTTGTCTTTTGTAGTTGAATTC 1841

OS Unidentified.
 OS Chimeric.
 PN WO200279239-A2.
 XX
 PD 10-OCT-2002.
 XX
 PF 31-JAN-2002; 2002WO-US003339.
 XX
 PR 31-JAN-2001; 2001US-0267522P.
 XX
 PA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
 XX
 PI Grogan CC, Hevey MC, Schmaljohn AL;
 XX
 DR WPI; 2003-040651/03.
 DR P-PSDB; ABJ18474.
 XX
 PT New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
 PT useful for inducing an immune response against infection of different
 PT filoviruses, specifically against both Ebola and Marburg viruses.
 XX
 PS Claim 16; Page 74-76; 94pp; English.
 XX
 CC The invention relates to a chimeric filovirus glycoprotein (GP) protein
 CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
 CC of GP2. The chimeric filovirus GP protein is useful for inducing an
 CC immune response against infection of different filoviruses, specifically
 CC against both Ebola and Marburg viruses by being used as a vaccine. This
 CC polynucleotide sequence represents a DNA encoding a chimeric filovirus
 CC protein of the invention
 XX
 SQ Sequence 2046 BP; 659 A; 517 C; 402 G; 468 T; 0 U; 0 Other;

Query Match 71.0%; Score 1308; DB 10; Length 2046;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAATCTCCCC 60
 DB 1 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACAAAAATCTCCCC 60

QY 61 ATTTAGAGATAGCTAGTAATCAACCCCAAAATGTGGATTCGGATGCTCCGGAAT 120
 DB 61 ATTTAGAGATAGCTAGTAATCAACCCCAAAATGTGGATTCGGATGCTCCGGAAT 120

QY 121 CTCGAGAGACAGAGAGCTCCATCTGATGGATTCACATGAGTGGGCAAAAAGTTGCT 180
 DB 121 CTCGAGAGACAGAGAGCTCCATCTGATGGATTCACATGAGTGGGCAAAAAGTTGCT 180

QY 181 GATTCCTCTTTGGAGGCATCCAAAGCGATGGGCTTTTCCAGGACAGGTGTACCTCCCAAGAA 240
 DB 181 GATTCCTCTTTGGAGGCATCCAAAGCGATGGGCTTTTCCAGGACAGGTGTACCTCCCAAGAA 240

QY 241 GTTGAATACAGAGAGGGGAGAGGCAAAACATGCTACATATAGTGTAAAGGATCCG 300
 DB 241 GTTGAATACAGAGAGGGGAGAGGCAAAACATGCTACATATAGTGTAAAGGATCCG 300

QY 301 TCTGGAAATCCTGCTGTAGATCCTCTCAACATCCGCTATCCGAAATGCAAA 360
 DB 301 TCTGGAAATCCTGCTGTAGATCCTCTCAACATCCGCTATCCGAAATGCAAA 360

QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
 DB 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420

QY 421 GGAGCATTTTTTCTGTATGATCGCATTCCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 480
 DB 421 GGAGCATTTTTTCTGTATGATCGCATTCCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 480

QY 481 ACTGAAGGACATAGCAGCTATGATGTCATATAGACAGTGCACAAAATGATTTTCTCG 540
 DB 481 ACTGAAGGACATAGCAGCTATGATGTCATATAGACAGTGCACAAAATGATTTTCTCG 540

QY 541 CGGCAAGGACAAAGGGTACCGTCAATATGAATCTGATCTTCTACTAATAAATATTGGACAAGT 600
 DB 541 CGGCAAGGACAAAGGGTACCGTCAATATGAATCTGATCTTCTACTAATAAATATTGGACAAGT 600

QY 601 AGTAAACGGAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGATACAAATTTCT 660
 DB 601 AGTAAACGGAACGCAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGATACAAATTTCT 660

QY 661 ACAAGGAACCAAAACATGTGTCTCCGTCCAAATATCTCCACACTGCCACACAGCCGCTCG 720
 DB 661 ACAAGGAACCAAAACATGTGTCTCCGTCCAAATATCTCCACACTGCCACACAGCCGCTCG 720

QY 721 GAGATCAAACTCAACAGCACCCCAACTGATGTCACCAAACTCAATATACCGAGACCAAGC 780
 DB 721 GAGATCAAACTCAACAGCACCCCAACTGATGTCACCAAACTCAATATACCGAGACCAAGC 780

QY 781 AGTCATGATGAGGACCTCGCAACATCGGCTCAGGGTCGGAGAACGAGAACCCACACA 840
 DB 781 AGTCATGATGAGGACCTCGCAACATCGGCTCAGGGTCGGAGAACGAGAACCCACACA 840

QY 841 ACTTCTGATGCGGTCAACAAAGGCGTTCATCAACAATGCCACCTCCCTCACCA 900
 DB 841 ACTTCTGATGCGGTCAACAAAGGCGTTCATCAACAATGCCACCTCCCTCACCA 900

QY 901 CAACCAAGCAGCCACACAGCAAGGAGGAAAACAAACCAATTCCTCAAGATGCTGTGACT 960
 DB 901 CAACCAAGCAGCCACACAGCAAGGAGGAAAACAAACCAATTCCTCAAGATGCTGTGACT 960

QY 961 GAATAGACAAAAATAACACAACTGCAACCGTCCATGCCCTCCATCAATACACTACCA 1020
 DB 961 GAATAGACAAAAATAACACAACTGCAACCGTCCATGCCCTCCATCAATACACTACCA 1020

QY 1021 ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTGCAACCAATTACAAAAC 1080
 DB 1021 ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTGCAACCAATTACAAAAC 1080

QY 1081 ACCACCAATGACACACAGAGCACAATCACTGAAAATGAGCAAAACAGTGCCTCCCTCG 1140
 DB 1081 ACCACCAATGACACACAGAGCACAATCACTGAAAATGAGCAAAACAGTGCCTCCCTCG 1140

QY 1141 ATACCAACCTGCTCCAAACGGAATCCCAACACAGCAAGGACCAAGCAGCAAAAAA 1200
 DB 1141 ATACCAACCTGCTCCAAACGGAATCCCAACACAGCAAGGACCAAGCAGCAAAAAA 1200

QY 1201 GGCCCGCCCAACAGGCAACCAACACGCAAAATGCCACAGCAAGGACCAAGCAGCAAAAAA 1260
 DB 1201 GGCCCGCCCAACAGGCAACCAACACGCAAAATGCCACAGCAAGGACCAAGCAGCAAAAAA 1260

QY 1261 CCCAGCTCGACTGCACAACTCTTTGTATATTTTCAAGAAAGCGATCG 1308
 DB 1261 CCCAGCTCGACTGCACAACTCTTTGTATATTTTCAAGAAAGCGATCG 1308

RESULT 3
 ACC57764
 ID ACC57764 standard; cDNA; 2103 BP.
 XX
 AC ACC57764;
 XX
 DT 28-JUL-2003 (first entry)
 XX
 DE Marburg virus envelope glycoprotein coding sequence.
 XX
 KW Lentivirus; pseudotyped virus; gene therapy; vector; gene; ss.
 XX
 OS Marburg virus.
 XX
 FT Key Location/Qualifiers
 FT CDS 10..2055
 FT /*tag= a
 FT /*product= "Envelope glycoprotein"

PN WO2003035849-A2.
 PD 01-MAY-2003.
 PP 28-OCT-2002; 2002WO-US034545.
 PP 26-OCT-2001; 2001US-0353221P.
 PR 26-OCT-2001; 2001US-0356436P.
 XX (MCCR/) MCCRAY P B.
 PA (SAND/) SANDERS D A.
 PA (DAVI/) DAVIDSON B L.
 PI McCreay PB, Sanders DA, Davidson BL;
 XX WPI: 2003-421416/39.
 DR P-PSDB; ABR42242.
 XX New pseudotyped lentivirus comprising a lentiviral capsid, a lipid bilayer and a Marburg glycoprotein disposed in the lipid bilayer, useful for eliciting an immune response against feline immunodeficiency virus infection.
 PS Disclosure; Page 36-39; 4lpp; English.
 XX The present sequence is the coding sequence for the envelope glycoprotein of Marburg virus. The invention provides methods for gene transfer to cells using glycoprotein-pseudotyped lentiviruses. The glycoprotein is preferably Marburg glycoprotein. The Marburg viral glycoprotein CC preferably has a mutation in the C-terminal portion (see ABR42243-47) that results in a higher titre production of the pseudotyped virus. The pseudotyped lentivirus is used to introduce nucleic acid sequences encoding a desired protein into a hepatocyte, brain glial or airway epithelial cell, in vitro or in vivo. The desired protein is preferably cystic fibrosis transmembrane conductance regulator, or is the low density lipoprotein receptor, alpha1-antitrypsin, ornithine transcarbamylase, Factor VIII or a high affinity glutamate receptor (all claimed)
 XX Sequence 2103 BP; 671 A; 533 C; 408 G; 491 T; 0 U; 0 Other;
 SQ
 Query Match 70.8%; Score 1303.4; DB 8; Length 2103;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGAGACACATGTTTCTTATCTAGTCTTATCTTAAATTCAGGGGACAAAATCTCC 60
 DB |||||
 QY 10 ATGAGACACATGTTTCTTATCTAGTCTTATCTTAAATTCAGGGGACAAAATCTCC 69
 DB |||||
 QY 61 ATTTTAGAGATAGTAAATATCAACCCCAAAATGTGGATTGGTATGCTCCGGAAT 120
 DB |||||
 QY 70 ATTTTAGAGATAGTAAATATCAACCCCAAAATGTGGATTGGTATGCTCCGGAAT 129
 DB |||||
 QY 121 CTCAGAGACAGAACGCTTCACTGTGATGGATTACACTGAGTGGGCAAAAATGTGCT 180
 DB |||||
 QY 130 CTCAGAGACAGAACGCTTCACTGTGATGGATTACACTGAGTGGGCAAAAATGTGCT 189
 DB |||||
 QY 181 GATTCCTCTTGGAGGATCCAGCGATGGCTTTTCAGGACAGGTTGACTCCCAAGAT 240
 DB |||||
 QY 190 GATTCCTCTTGGAGGATCCAGCGATGGCTTTTCAGGACAGGTTGACTCCCAAGAT 249
 DB |||||
 QY 241 GTTGAGTACACAGAGGGGAGGAGCAAAAATGCTACATATAGTGTAAAGGATCC 300
 DB |||||
 QY 250 GTTGAGTACACAGAGGGGAGGAGCAAAAATGCTACATATAGTGTAAAGGATCC 309
 DB |||||
 QY 301 TCTGGAATAATCTTGTGTGTAGATCTCTTAACAACATCCGTGACTATCCGAAATGC 360
 DB |||||
 QY 310 TCTGGAATAATCTTGTGTGTAGATCTCTTAACAACATCCGTGACTATCCGAAATGC 369
 DB |||||
 QY 361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGATCGCCCTTCATTTATGG 420
 DB |||||
 QY 370 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGATCGCCCTTCATTTATGG 429
 DB |||||

QY 421 GGAGCATTTTTTCTGTATGATCGCATTCCTCCACAAATGTATACCGAGGAAAATCTTTC 480
 DB |||||
 QY 430 GGAGCATTTTTTCTGTATGATCGCATTCCTCCACAAATGTATACCGAGGAAAATCTTTC 489
 DB |||||
 QY 481 ACTGAGGGAACATAGCAGCTATGATTTGTCATAGACAGTGCACAAAATGATTTTCTCG 540
 DB |||||
 QY 490 ACTGAGGGAACATAGCAGCTATGATTTGTCATAGACAGTGCACAAAATGATTTTCTCG 549
 DB |||||
 QY 541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTATAAATATTGGACAAAGT 600
 DB |||||
 QY 550 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTATAAATATTGGACAAAGT 609
 DB |||||
 QY 601 AGTAAACGGAACGCAAAACGAATGACATGATGTTTTCGGGCTCTTCAAGAAATACAAATCT 660
 DB |||||
 QY 610 AGTAAACGGAACGCAAAACGAATGACATGATGTTTTCGGGCTCTTCAAGAAATACAAATCT 669
 DB |||||
 QY 661 ACAAGAACCACCAATGTGCTCCGTCGCAAAATACCTCCACACCTGCCACAGCCCGTCGG 720
 DB |||||
 QY 670 ACAAGAACCACCAATGTGCTCCGTCGCAAAATACCTCCACACCTGCCACAGCCCGTCGG 729
 DB |||||
 QY 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCATACCGGACCCCAAGC 780
 DB |||||
 QY 730 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCATACCGGACCCCAAGC 789
 DB |||||
 QY 781 AGTATGATGAGGACCTTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 840
 DB |||||
 QY 790 AGTATGATGAGGACCTTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 849
 DB |||||
 QY 841 ACTTCTGATGGTTCACCAAGCAGGCTTTTCATCAAAATGCCACCTCCCTCCACCA 900
 DB |||||
 QY 850 ACTTCTGATGGTTCACCAAGCAGGCTTTTCATCAAAATGCCACCTCCCTCCACCA 909
 DB |||||
 QY 901 CAACCAACGACGCGCACACAGAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 960
 DB |||||
 QY 910 CAACCAACGACGCGCACACAGAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 969
 DB |||||
 QY 961 GAATCTAGACAAAATAAACAACATGACAACTCCGTCATGCCCCCTCATATAACTACCA 1020
 DB |||||
 QY 970 GAATCTAGACAAAATAAACAACATGACAACTCCGTCATGCCCCCTCATATAACTACCA 1029
 DB |||||
 QY 1021 ATCTCTACTACACACCTCCCAACACAACTTCAGACTCTCTCTGCACCATTCACAAAAC 1080
 DB |||||
 QY 1030 ATCTCTACTACACACCTCCCAACACAACTTCAGACTCTCTCTGCACCATTCACAAAAC 1089
 DB |||||
 QY 1081 ACCCAATATGACACACACAGACACAACTCACTGAAAAATGAGCAAAACCAAGTCCCTCTCG 1140
 DB |||||
 QY 1090 ACCCAATATGACACACACAGACACAACTCACTGAAAAATGAGCAAAACCAAGTCCCTCTCG 1149
 DB |||||
 QY 1141 ATAACAACCTCTGCTCCAAACGCGGAAATCCCAACACAGCAAAAGAGCACCGAGGCAAAA 1200
 DB |||||
 QY 1150 ATAACAACCTCTGCTCCAAACGCGGAAATCCCAACACAGCAAAAGAGCACCGAGGCAAAA 1209
 DB |||||
 QY 1201 GGCCCCGACACACGCGGCAAAACAGCAAAATGAGCAATTTTCCACAGTCTCCCTCCCTCC 1260
 DB |||||
 QY 1210 GGCCCCGACACACGCGGCAAAACAGCAAAATGAGCAATTTTCCACAGTCTCCCTCCCTCC 1269
 DB |||||
 QY 1261 CCAGCTGACTGCAACAACTCTGTATATTTTCAAGAAAGCGA 1305
 DB |||||
 QY 1270 CCAGCTGACTGCAACAACTCTGTATATTTTCAAGAAAGCGA 1314
 DB |||||

RESULT 4
 AAD04042
 ID AAD04042 standard; DNA; 2164 BP.
 XX AAD04042;
 AC AAD04042;
 XX
 DT 02-JUL-2001 (first entry)
 XX
 DE Marburg virus Musoke strain glycoprotein (GP) DNA.
 XX Glycoprotein; GP; immune response; vaccine; antiviral;
 KW type I transmembrane protein; ds.

[illegible]

DE VRC6702 (pVR1012-x/s Marburg GP (dTM)) plasmid.
 XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;
 KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
 KW viral infection; filovirus; circular; cyclic; ds.
 XX Synthetic.
 OS WO2003028632-A2.
 PN 10-APR-2003.
 PD 24-SEP-2002; 2002WO-US030251.
 XX 01-OCT-2001; 2001US-0326476P.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Nabel GU, Yang Z, Sullivan N, Sanchez A;
 PI WPI; 2003-371961/35.
 DR New bimodal priming and boosting compositions, useful as viral vaccines,
 XX specifically for eliciting an immune response against a filovirus or a
 PT disease caused by infection with filovirus.
 PT Claim 1; Page 188-190; 219pp; English.
 PS The present invention relates to a bimodal priming composition and
 XX boosting composition for priming and boosting an immune response to an
 CC antigen in an individual. The compositions comprise (a) a priming
 CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
 CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
 CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
 CC DNA plasmid selected from AC71521-AC71563 and (b) a boosting
 CC composition comprised of a replication-deficient adenovirus, comprising a
 CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
 CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
 CC epitope-bearing domain, or a replication deficient adenovirus selected
 CC from constructs AC71521-AC71563. The compositions are useful as viral
 CC vaccines, specifically for eliciting an immune response against a
 CC filovirus or a disease caused by infection with filovirus e.g. Ebola
 CC virus infection. This vector expresses the Marburg virus glycoprotein
 CC without its transmembrane and intracellular domains
 XX Sequence 7005 BP; 1931 A; 1758 G; 1588 G; 1728 T; 0 U; 0 Other;
 SQ

Query Match 70.8%; Score 1303.4; DB 10; Length 7005;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
 DB 2034 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 2093

QY 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
 DB 2094 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 2153

QY 121 CTCGAGAGACAGAGACGTCATCTGATGGGATTACACTGAGTGGGCAAAAAGTTGCT 180
 DB 2154 CTCGAGAGACAGAGACGTCATCTGATGGGATTACACTGAGTGGGCAAAAAGTTGCT 2213

QY 181 GATTCCTCTTGGAGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
 DB 2214 GATTCCTCTTGGAGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 2273

QY 241 GTTGTAGTACAGAGGGGAGGAGCAAAAATGCTACAATATAGTGAACGGATCCC 300
 DB 2274 GTTGTAGTACAGAGGGGAGGAGCAAAAATGCTACAATATAGTGAACGGATCCC 2333

QY 301 TCTGGAAAAATCTTGCTGTAGATCTCTTAACAACATCCGTGACTATCCGAAATGCAAA 360
 DB 301 TCTGGAAAAATCTTGCTGTAGATCTCTTAACAACATCCGTGACTATCCGAAATGCAAA 360

DB 2334 TCTGGAAAAATCTTGCTGTAGATCTCTTAACAACATCCGTGACTATCCGAAATGCAAA 2393

QY 361 ACTATCCATCATATTCAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
 DB 2394 ACTATCCATCATATTCAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATGG 2453

QY 421 GGAGCATTTTCTGTATGATCGCATTTGCCCTCCACAAATGTACCGAGGCAAAATCTTTC 480
 DB 2454 GGAGCATTTTCTGTATGATCGCATTTGCCCTCCACAAATGTACCGAGGCAAAATCTTTC 2513

QY 481 ACTGAAGGGACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCTCG 540
 DB 2514 ACTGAAGGGACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCTCG 2573

QY 541 CGGCAAGGACCAAGGTACCGTATATGAATCTGACTTCTACTATAAATATTTGGACAAGT 600
 DB 2574 CGGCAAGGACCAAGGTACCGTATATGAATCTGACTTCTACTATAAATATTTGGACAAGT 2633

QY 601 AGTAACGGGAACGCAAAACGAATGACACTGGATGTTTGGGGCTCTTTCAAGAATACAAATCT 660
 DB 2634 AGTAACGGGAACGCAAAACGAATGACACTGGATGTTTGGGGCTCTTTCAAGAATACAAATCT 2693

QY 661 ACAAGACCAACCAATGCTCCGTCCAAAATACCTCCACCACTGCCCCACAGCCCGTCCG 720
 DB 2694 ACAAGACCAACCAATGCTCCGTCCAAAATACCTCCACCACTGCCCCACAGCCCGTCCG 2753

QY 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCCAACCAACTCAATACCGACCCCAAGC 780
 DB 2754 GAGATCAAACTCACAAAGCACCCCAACTGATGCCAACCAACTCAATACCGACCCCAAGC 2813

QY 781 AGTGATGATGAGGACCTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
 DB 2814 AGTGATGATGAGGACCTCGCAACATTCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 2873

QY 841 ACTTCTGATGGGTCCACCAACGCAAGGGCTTTTCATCAACAATGCCACCCCACTCCCTCACA 900
 DB 2874 ACTTCTGATGGGTCCACCAACGCAAGGGCTTTTCATCAACAATGCCACCCCACTCCCTCACA 2933

QY 901 CAACCAAGCACGCCACACAGAGGAGGAAACAAACAAACCACTTCCCAAGATGCTGTGACT 960
 DB 2934 CAACCAAGCACGCCACACAGAGGAGGAAACAAACAAACCACTTCCCAAGATGCTGTGACT 2993

QY 961 GAATCTAGACAAAAATAACAACTGACAAACCGTCCATGCCCCCTCATATAACCTACCA 1020
 DB 2994 GAATCTAGACAAAAATAACAACTGACAAACCGTCCATGCCCCCTCATATAACCTACCA 3053

QY 1021 ATCTCTACTACAAACCTCCAAACACAACTTCAGCACTCTCTGACCACTTACCAAAAC 1080
 DB 3054 ATCTCTACTACAAACCTCCAAACACAACTTCAGCACTCTCTGACCACTTACCAAAAC 3113

QY 1081 ACCACCAATGACACACACAGAGCACAATCACTGAAAAATGAGCAAAACAGTGCCCCCTCG 1140
 DB 3114 ACCACCAATGACACACACAGAGCACAATCACTGAAAAATGAGCAAAACAGTGCCCCCTCG 3173

QY 1141 ATAAACAACTTGCCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAGCAGAGCAAAAAA 1200
 DB 3174 ATAAACAACTTGCCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAGCAGAGCAAAAAA 3233

QY 1201 GGCCCCGCCACACAGGACCAACACGACCAATGAGCATTTTCCAGCTCCTCCCCCACC 1260
 DB 3234 GGCCCCGCCACACAGGACCAACACGACCAATGAGCATTTTCCAGCTCCTCCCCCACC 3293

QY 1261 CCCAGCTCGATGCAACACATCTTGTATATTTTCAAGAAGAAAGCGA 1305
 DB 3294 CCCAGCTCGATGCAACACATCTTGTATATTTTCAAGAAGAAAGCGA 3338

RESULT 6
 ACC71550
 ID ACC71550 standard; DNA; 7778 BP.
 XX AC ACC71550;
 XX

DT	10-JUL-2003	(first entry)
XX	VR6701	(pVR1012-Marburg) plasmid.
XX		
KW	Virucide;	vaccine; immune response; Ebola virus; Marburg virus;
KW	Lassa virus;	retrovirus; paramyxovirus; influenza virus; adenovirus;
KW	viral infection;	filovirus; circular; cyclic; ds.
XX		
XX	Synthetic.	
XX	WO2003028632-A2.	
PN		
XX	10-APR-2003.	
PD		
XX	24-SEP-2002;	2002WO-US030251.
XX		
XX	01-OCT-2001;	2001US-0326476P.
PR		
XX	(USSH)	US DEPT HEALTH & HUMAN SERVICES.
PA		
XX	Nabel GJ,	Yang Z, Sullivan N, Sanchez A;
PI		
XX	WPI;	2003-371961/35.
DR		
XX		
PT	New bimodal	priming and boosting compositions, useful as viral vaccines,
PT	specifically	for eliciting an immune response against a filovirus or a
PT	disease	caused by infection with filovirus.
XX		
PS	Claim 1;	Page 185-187; 219pp; English.
XX		
CC	The present	invention relates to a bimodal priming composition and
CC	boosting	composition for priming and boosting an immune response to an
CC	antigen	in an individual. The compositions comprise (a) a priming
CC	composition	comprised of a DNA plasmid comprising a nucleic acid molecule
CC	encoding	Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC	virus	glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC	DNA	plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC	composition	comprised of a replication-deficient adenovirus, comprising a
CC	nucleic	acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC	paramyxovirus,	or influenza virus glycoprotein or nucleoprotein or
CC	epitope-	bearing domain, or a replication deficient adenovirus selected
CC	from	constructs ACC71521-ACC71563. The compositions are useful as viral
CC	vaccines,	specifically for eliciting an immune response against a
CC	filovirus	or a disease caused by infection with filovirus e.g. Ebola
CC	virus	infection. The present sequence has a Marburg glycoprotein (GP)
CC	open	reading frame, Musoke strain, cloned into a VRC6700 backbone
XX		
SQ	Sequence	7778 BP; 2148 A; 1892 C; 1719 G; 2019 T; 0 U; 0 Other;
	Query	Match
	Best	Local Similarity 70.8%; Score 1303.4; DB 10; Length 7778;
	Matches	1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy	1	ATGAAGACCAATGTTTCCTTATCAAGTCTTATCTTAATTCAGGACAAAATCTCCCC 60
Db	2034	ATGAAGACCAATGTTTCCTTATCAAGTCTTATCTTAATTCAGGACAAAATCTCCCC 2093
Qy	61	ATTTTACGATAGCTAGTATATCAACCCCAAAATGCGATTCCTCGGNACT 120
Db	2094	ATTTTACGATAGCTAGTATATCAACCCCAAAATGCGATTCCTCGGNACT 2153
Qy	121	CTCCAGAAGACAGAACGTCCTCATCTCATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 180
Db	2154	CTCCAGAAGACAGAACGTCCTCATCTCATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 2213
Qy	181	GATTCCCTTTGGAGGCATCCACCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db	2214	GATTCCCTTTGGAGGCATCCACCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 2273
Qy	241	GTTGAGTACACAGAGGGGGAGGACCAAAACATGCTACAAATTAAGTGTAAAGGATCCC 300
Db	2274	GTTGAGTACACAGAGGGGGAGGAGCCAAAACATGCTACAAATTAAGTGTAAAGGATCCC 2333

Qy	301	TCTGGAATAATCCCTTGCTGTAGATCTCTCTTACCAACATCCGTGACTATCGAAATGCAAA	360
Db	2334	TCTGGAATAATCCCTTGCTGTAGATCTCTCTTACCAACATCCGTGACTATCTTAAATGCAAA	2393
Qy	361	ACTATCCATCATATTCGAAGTCAAAACCCCTCATGCAACAGGGATCGCCCTTCATTTATGG	420
Db	2394	ACTATCCATCATATTCGAAGTCAAAACCCCTCATGCAACAGGGATCGCCCTTCATTTATGG	2453
Qy	421	GGAGCATTTTTTCTGTATGATCGCATTTGCTCCACAACAATGTATCGAGGCAAGTCTTTC	480
Db	2454	GGAGCATTTTTTCTGTATGATCGCATTTGCTCCACAACAATGTATCGAGGCAAGTCTTTC	2513
Qy	481	ACTGAAGGGAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTTTCTCG	540
Db	2514	ACTGAAGGGAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTTTCTCG	2573
Qy	541	CGGCAAGGACAAAGGTACCGTCAATATGAATCTGCATTTCTACTAATAAATATGGACAAGT	600
Db	2574	CGGCAAGGACAAAGGTACCGTCAATATGAATCTGCATTTCTACTAATAAATATGGACAAGT	2633
Qy	601	AGTAAACGGAACGGAAACGAATGACATCTGGATGTTTGGCGCTCTTCAAGAATATCAATTTCT	660
Db	2634	AGTAAACGGAACGGAAACGAATGACATCTGGATGTTTGGCGCTCTTCAAGAATATCAATTTCT	2693
Qy	661	ACAAGAACCAACATATGTCTCGTCCCAATAACCTCCACCATGCCCCACAGCCGCTCCG	720
Db	2694	ACAAGAACCAACATATGTCTCGTCCCAATAACCTCCACCATGCCCCACAGCCGCTCCG	2753
Qy	721	GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAACTCAATACACGGACCCCAAGC	780
Db	2754	GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAACTCAATACACGGACCCCAAGC	2813
Qy	781	AGTGATGATGAGACCTCGCAACATCGGGCTCAGGGTCCGGAGAACGAGAACCCCAACA	840
Db	2814	AGTGATGATGAGACCTCGCAACATCGGGCTCAGGGTCCGGAGAACGAGAACCCCAACA	2873
Qy	841	ACTTCTGATCGGCTCAACAGCAAGGCTTTTCAATCAACAAATGACCCACTCTCCTCAACA	900
Db	2874	ACTTCTGATCGGCTCAACAGCAAGGCTTTTCAATCAACAAATGACCCACTCTCCTCAACA	2933
Qy	901	CAACCAAGCACGCGCACAGCAAGGAGGAAACAAACAAACCAATTTCCCAAGATGCTGTGACT	960
Db	2934	CAACCAAGCACGCGCACAGCAAGGAGGAAACAAACAAACCAATTTCCCAAGATGCTGTGACT	2993
Qy	961	GAACTAGACAAAATAACACAACTGACACACCGTCCATGCCCCCTCATATACACTACCACA	1020
Db	2994	GAACTAGACAAAATAACACAACTGACACACCGTCCATGCCCCCTCATATACACTACCACA	3053
Qy	1021	ATCTCTACTAAACAAACCTCCAAACCAACTTCAGCACTCTCTGTGCACCATTTACAAAAC	1080
Db	3054	ATCTCTACTAAACAAACCTCCAAACCAACTTCAGCACTCTCTGTGCACCATTTACAAAAC	3113
Qy	1081	ACCAACCAATGACAACACACAGAGCAAAATCACTGAAAATGAGCAAAACCAAGTGCCTCTCG	1140
Db	3114	ACCAACCAATGACAACACACAGAGCAAAATCACTGAAAATGAGCAAAACCAAGTGCCTCTCG	3173
Qy	1141	ATAACAAACCTGCTCCACGGGAATCCACACACAGCAAGAGACCAAGTGCCTCTCTCTCT	1200
Db	3174	ATAACAAACCTGCTCCACGGGAATCCACACACAGCAAGAGACCAAGTGCCTCTCTCTCT	3233
Qy	1201	GGCCCCGCCACAAACGGCACCAACACAGCAAAATGAGCATTTTCAACAGTCTCTCCCCCACC	1260
Db	3234	GGCCCCGCCACAAACGGCACCAACACAGCAAAATGAGCATTTTCAACAGTCTCTCCCCCACC	3293
Qy	1261	CCGAGCTCGATCGCAACAACTTTGTATATTTTCAAGAAAGGCA	1305
Db	3294	CCGAGCTCGATCGCAACAACTTTGTATATTTTCAAGAAAGGCA	3338

RESULT 7
ACC71552
ID ACC
XX


```
AC ACCT1552;
XX
DT 10-JUL-2003 (first entry)
DE
DE VRC6710 (pAdapt Marburg GP(dtm)) plasmid.
XX
XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX
OS Synthetic.
XX
XX WO2003028632-A2.
PN
XX
XX 10-APR-2003.
PD
XX
XX 24-SEP-2002; 2002WO-US030251.
XX
XX 01-OCT-2001; 2001US-0326476P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX
XX WPI; 2003-371961/35.
DR
XX
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
XX Claim 1; Page 190-193; 219pp; English.
XX
XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. The present sequence is an adenovirus shuttle vector,
CC expressing the Marburg virus glycoprotein without its transmembrane and
CC intracellular domains
XX
XX Sequence 8256 BP; 2048 A; 2049 C; 2108 G; 2051 T; 0 U; 0 Other;
SQ
Query Match 70.8%; Score 1303.4; DB 10; Length 8256;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
DB 1431 ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 1490
QY 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGTCCTCGAACT 120
DB 1491 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGTCCTCGAACT 1550
QY 121 CTCCAGACACAGACAGCTCTCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTCCT 180
DB 1551 CTCCAGACACAGACAGCTCTCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTCCT 1610
QY 181 GATTCCTCTTTGGAGGATCCCAAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAAAT 240
DB 1611 GATTCCTCTTTGGAGGATCCCAAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAAAT 1670
QY 241 GTTGAGTACACAGAGGGGGAGGAGCCAAAAACATGCTCAATATAAGTGTAAACGGATCCC 300
```

RESULT 8

AAZ87211
ID AAZ87211 standard; cDNA; 11460 BP.
XX
AC AAZ87211;
XX
15-SEP-2003 (revised)
DT
08-MAY-2000 (first entry)
XX
VEE replicon comprising Marburg virus genes.
DE
XX
VEE virus replicon; Venezuelan equine encephalitis; Marburg virus; MBGV;
KW filovirus; glycoprotein; GP; nucleoprotein; NP; structural protein; VP35;
KW VP35; VP30; VP24; genetic vaccine; antigen delivery;
KW Marburg haemorrhagic fever; cyclic; circular; ss.
XX
Marburg virus; str. Musoke.
OS Venezuelan equine encephalitis virus.
OS Chimeric.
XX
FH Location/Qualifiers
CDS 104..2182
FT /*tag= a
FT /product= "Marburg virus nucleoprotein (NP, AAY77128)"
FT CDS 2944..3933
FT /*tag= b
FT /product= "Marburg virus structural protein VP35
FT (AAY77130)"
FT CDS 4567..5478
FT /*tag= c
FT /product= "Marburg virus structural protein VP40
FT (AAY77129)"
FT CDS 5940..7985
FT /*tag= d
FT /product= "Marburg virus glycoprotein (GP, AAY77127)"
FT CDS 8864..9697
FT /*tag= e
FT /product= "Marburg virus structural protein VP30
FT (AAY77131)"
FT CDS 10200..10961
FT /*tag= f
FT /product= "Marburg virus structural protein VP24
FT (AAY77132)"
XX
WO200000616-A2.
XX
06-JAN-2000.
XX
21-JUN-1999; 99WO-US014174.
XX
29-JUN-1998;. 98US-0091403P.
XX (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX Hevey MC, Negley DL, Pushko P, Smith JF, Schmaljohn AL;
PI
XX WPI; 2000-160676/14.
DR P-PSDB; AAY77127, AAY77128, AAY77129, AAY77130, AAY77131, AAY77132,
DR AAY77133.
XX
XX Novel Marburg virus vaccines used to induce an immune response against
PT the infection in nonhuman primates.
XX
XX Claim 1; Page 56-57; 57pp; English.
XX
XX The invention relates to novel Marburg virus (MBGV) vaccines. The vaccine
CC of the invention comprises a Venezuelan equine encephalitis virus (VEE)
CC replicon containing copies of the genes encoding Marburg virus (Musoke
CC strain) glycoprotein (GP), a deletion mutant GP (GP-Delta-TM),
CC nucleoprotein (NP), and structural proteins VP40, VP35, VP30 and VP24.
CC The Marburg virus genes replace the VEE virus structural protein genes;
CC the result is a self-replicating RNA molecule that encodes its own
CC replicase and transcriptase functions, and in addition makes abundant
CC quantities of the Marburg virus proteins. When replicon RNA is

transfected into eukaryotic cells, along with two helper RNAs that
express the VEE virus structural proteins, the replicon RNA is packaged
into VEE virus-like particles by the VEE virus structural proteins, which
are provided in trans. Since the helper RNAs lack packaging signals
necessary for further propagation, the resulting VEE replicon particles
(VRPs) which are produced are infectious for one cycle but are defective
thereafter. On VRP infection of a cell, an abortive infection occurs
whereby the cell produces the Marburg virus proteins, is ultimately
killed by the infection, but does not produce any viral progeny. The VEE
replicon provides a potent tool for vaccination with Marburg virus
antigens. The replicons, vectors and constructs are used to produce
vaccines against Marburg virus (MBGV) infection (Marburg haemorrhagic
fever) in mammals, to elicit immune responses against Marburg antigens,
to confer protective immunity, and to reduce disease symptoms and reduce
the severity of disease. Studies of non-human primates vaccinated with
the replicon indicate that the vaccine will be efficient in protecting
humans against Marburg virus. Prior art Marburg virus vaccination
strategies have used formalin-inactivated Marburg virus, which are only
partially successful at protecting against Marburg virus infection. The
present sequence represents a cDNA corresponding to a VEE virus replicon
encoding Marburg virus GP, GP-delta-TM, NP, VP40, VP35, VP30 and VP24.
CC (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 11460 BP; 3674 A; 2382 C; 2145 G; 3259 T; 0 U; 0 Other;
Query Match 70.8%; Score 1303.4; DB 3; Length 11460;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTTCAAGGACAAAAAATCTCCCC 60
DB 5940 ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTTCAAGGACAAAAAATCTCCCC 5999
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTTCGGTATGCTCCGGAAT 120
DB 6000 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTTCGGTATGCTCCGGAAT 6059
QY 121 CTCAGAAAGACAGAAGCGTCCATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT 180
DB 6060 CTCAGAAAGACAGAAGCGTCCATCTGATGGGATTTCACACTGAGTGGGCAAAAAGTTGCT 6119
QY 181 GATTCCCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGACAGAGGTGTACCTCCCAAGAAAT 240
DB 6120 GATTCCCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGACAGAGGTGTACCTCCCAAGAAAT 6179
QY 241 GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATATAAGTGAACGATCCC 300
DB 6180 GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATATAAGTGAACGATCCC 6239
QY 301 TCTGAAAAATCCTTGCTGTTAGATCCTCTTACCAACATCCTGTGACTATCCGAAATGCAAA 360
DB 6240 TCTGAAAAATCCTTGCTGTTAGATCCTCTTACCAACATCCTGTGACTATCCGAAATGCAAA 6299
QY 361 ACTATCCATCATATTCAAGGTCAAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
DB 6300 ACTATCCATCATATTCAAGGTCAAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 6359
QY 421 GGAGCATTTTTTCTGTATGATCGCATTTGCCCTCCACAAATGTATCCGAGGCAAAAGTCTTC 480
DB 6360 GGAGCATTTTTTCTGTATGATCGCATTTGCCCTCCACAAATGTATCCGAGGCAAAAGTCTTC 6419
QY 481 ACTGAAGGGAACATAGCAGCTTATGATTGTCAATAGACAGTGCACAAAAATGATTTCTCG 540
DB 6420 ACTGAAGGGAACATAGCAGCTTATGATTGTCAATAGACAGTGCACAAAAATGATTTCTCG 6479
QY 541 CGGCAGACAGAGGTACCGTACATATGATCTGACTTCTACTATAATAATATGGCAAGT 600
DB 6480 CGGCAGACAGAGGTACCGTACATATGATCTGACTTCTACTATAATAATATGGCAAGT 6539
QY 601 AGTAACGGAACGCAAAACGAATGACACTGGATGTTTGGCGGCTCTTCAAGAATACAAATCT 660
DB 6540 AGTAACGGAACGCAAAACGAATGACACTGGATGTTTGGCGGCTCTTCAAGAATACAAATCT 6599

```

QY 661 ACAAGAACCAAAATGCTCGTCCAAATAATCCTCCACCACTGCCACAGCCCGTCCG 720
Db 6600 ACAAGAACCAAAATGCTCGTCCAAATAATCCTCCACCACTGCCACAGCCCGTCCG 6659
QY 721 GAGATCAACTCACAGACCCCACTGATGCGGAGAACTCAATACCAAGCCCAAGC 780
Db 6660 GAGATCAACTCACAGACCCCACTGATGCGGAGAACTCAATACCAAGCCCAAGC 6719
QY 781 AGTGATGATGAGGACCTCCCAACATCCGGCTCAGGGTCGGAGAAACGAAACCCACACA 840
Db 6720 AGTGATGATGAGGACCTCCCAACATCCGGCTCAGGGTCGGAGAAACGAAACCCACACA 6779
QY 841 ACTTCTGATGCGGTCCACCAAGAGGCTTTTATCAACAATGCCACCACTCCCTCACA 900
Db 6780 ACTTCTGATGCGGTCCACCAAGAGGCTTTTATCAACAATGCCACCACTCCCTCACA 6839
QY 901 CAACCAAGCACGCCACAGCAAGGAGGAAACACACACCAATCCCAAGATGCTGTGACT 960
Db 6840 CAACCAAGCACGCCACAGCAAGGAGGAAACACACACCAATCCCAAGATGCTGTGACT 6899
QY 961 GAACTAGACAAAATAACAACTGACAAACCGTCCATGCCCTCATAACTACCACTACCACA 1020
Db 6900 GAACTAGACAAAATAACAACTGACAAACCGTCCATGCCCTCATAACTACCACTACCACA 6959
QY 1021 ATCTCTACTAACAACTCCAAACAACTTTCAGCACTCTCTCTGCACCACTTACAAAAC 1080
Db 6960 ATCTCTACTAACAACTCCAAACAACTTTCAGCACTCTCTCTGCACCACTTACAAAAC 7019
QY 1081 ACCACCAATGACACACACAGACACATCACTGAAATGACAAACCACTGCTCCCTCTG 1140
Db 7020 ACCACCAATGACACACAGACACATCACTGAAATGACAAACCACTGCTCCCTCTG 7079
QY 1141 ATAAACAACTGCTCTCCAAACGGGAAATCCACACAGCAAGAGCACAGCAAGCAAAA 1200
Db 7080 ATAAACAACTGCTCTCCAAACGGGAAATCCACACAGCAAGAGCACAGCAAGCAAAA 7139
QY 1201 GGCCCCGCCACAAACGGGACCAACAGCAAAATGAGCAATTTCCACCACTCTCTCCCCCA 1260
Db 7140 GGCCCCGCCACAAACGGGACCAACAGCAAAATGAGCAATTTCCACCACTCTCTCCCCCA 7199
QY 1261 CCAGCTGCACTGCACACACATCTTGATATTTTCAGNAGAAAGCGA 1305
Db 7200 CCAGCTGCACTGCACACACATCTTGATATTTTCAGNAGAAAGCGA 7244

```

RESULT 9

AZ51039
ID AZ51039 standard; DNA; 2100 BP.

AC AZ51039;

DT 05-JUN-2000 (first entry)

DE Marburg virus envelope glycoprotein DNA.

KW Marburg virus envelope glycoprotein; viral glycoprotein;
KW pseudotyped retrovirus; MLV; Moloney murine leukemia virus; antibody;
KW transduction; screening agent; immunological agent;
KW pharmacological agent; ss.

OS Marburg virus.

EH Key Location/Qualifiers

FT CDS 10.2052

FT /*tag= a

FT /product= "Marburg virus envelope glycoprotein"

PN W0200008131-A2.

PD 17-FEB-2000.

XX 04-AUG-1999; 99MO-US017702.

XX

PR 04-AUG-1998; 98US-0095242P.
PR 15-DEC-1998; 98US-0112405P.
XX
PA (PURD) PURDUE RES FOUND.

XX Sanders DA, Kuhn RJ, Jeffers SA, Sharkey CM, North CL;
PI Fischbach MA;
XX

XX WPI; 2000-224030/19.
DR P-PSDB; AAV70075.

XX Cells that produce inventive pseudotyped retroviruses having a broad host
PT range useful for introducing nucleotide sequences into target cells.

XX Example 11; Page 60-61; 65pp; English.

XX The present sequence encodes the Marburg virus envelope glycoprotein.
CC Plasmid pMBGP1 was produced by cloning into the plasmid pSP72 nucleotide
CC sequences corresponding to nucleotides 5931-8033 from the Marburg virus
CC genome. This is used for generation of stable cell lines transiently
CC producing Marburg-MLV (Moloney murine leukemia virus) pseudotyped
CC retrovirus with a broad host range. These cells having different viral
CC glycoproteins in its lipid bilayer with a pseudotyped retrovirus is
CC transduced with a desired ribonucleotide sequence. This is used to
CC identify screening agents effective in blocking viral entry into a cell.
CC These agents may be immunological agents like monoclonal or polyclonal
CC antibodies. The pharmacological agents include proteins, peptides or
CC various chemical agents. The pseudotyped retrovirus may be useful in
CC methods of identifying cell surface receptors that allow viral entry

XX SQ Sequence 2100 BP; 670 A; 532 C; 408 G; 490 T; 0 U; 0 Other;

Query Match 69.9%; Score 1287.4; DB 3; Length 2100;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1301; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 ATGAAGACCAACATGTTCTTATCAGTCTTATCTTTAATCAAGGACAAAAATCTCCCC 60

Db 10 ATGAAGACCAACATGTTCTTATCAGTCTTATCTTTAATCAAGGACAAAAATCTCCCC 69

QY 61 ATTTTAGAGTAGTAGTAAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120

Db 70 ATTTTAGAGTAGTAGTAAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 129

QY 121 CTCAGAAAGACAGAGACGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180

Db 130 CTCAGAAAGACAGAGACGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 189

QY 181 GATTCCCTTTGGAGGCAATCCAAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAA 240

Db 190 GATTCCCTTTGGAGGCAATCCAAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAA 249

QY 241 GTTGAGTACAGAGGGGGGAGGAGCCAAACATCTACATATAGTGTACCGATCC 300

Db 250 GTTGAGTACAGAGGGGGGAGGAGCCAAACATCTACATATAGTGTACCGATCC 309

QY 301 TCTGAAAAATCTCTGCTGTAGATCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360

Db 310 TCTGAAAAATCTCTGCTGTAGATCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 369

QY 361 ACTATCCATCATATTCAAAGGTCAAAACCTCATGACAGGGGATCCGCTTCATTTATGG 420

Db 370 ACTATCCATCATATTCAAAGGTCAAAACCTCATGACAGGGGATCCGCTTCATTTATGG 429

QY 421 GGAGCATTTTTTCTGTATGATGCGATTGCTCCCAACAAATGTACGGAGCAAGCTTTC 480

Db 430 GGAGCATTTTTTCTGTATGATGCGATTGCTCCCAACAAATGTACGGAGCAAGCTTTC 489

QY 481 ACTGAAGGGGAACATAGCAGCTATGATTGTCAATAAGACAGTGCACAAAATGATTTCTCG 540

Db 490 ---GAAGGGACATAGCAGCTATGATTGTCAATAAGACAGTGCACAAAATGATTTCTCG 546

QY 541 CGGCAAGGACAGGGTACCGTCATATGAATCTGACTTCTAATAAATATTTGGACAAGT 600

```
Db 547 CGCAGGAGCAGGAGTACCGTCTATGATCTGACTTCTACTAATAATATGACAGT 606
Qy 601 AGTAAACGGAACGCAACGAAATGACACTGATGTTTGGCGCTCTTCAAGAAATACAAATCT 660
Db 607 AGTAAACGGAACGCAACGAAATGACACTGATGTTTGGCGCTCTTCAAGAAATACAAATCT 666
Qy 661 ACAGAGACCAACCAATGCTCGGTCCAAATACCTCCACCACTGCCACAGCCCGTGG 720
Db 667 ACAGAGACCAACCAATGCTCGGTCCAAATACCTCCACCACTGCCACAGCCCGTGG 726
Qy 721 GAGATCAAACTCACAAGCACCCCAACTGATGCGCACCAAACTCAATACCAAGCACCAAGC 780
Db 727 GAGATCAAACTCACAAGCACCCCAACTGATGCGCACCAAACTCAATACCAAGCACCAAGC 786
Qy 781 AGTGATGATGAGGACCTCGCAACATCCCGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
Db 787 AGTGATGATGAGGACCTCGCAACATCCCGCTCAGGGTCCGGAGAACGAGAACCCCAACA 846
Qy 841 ACTTCTGATGCGGTACCAAGCAAGGCTTTTCAATCAAAATGCGCACCCACTCCCTCACA 900
Db 847 ACTTCTGATGCGGTACCAAGCAAGGCTTTTCAATCAAAATGCGCACCCACTCCCTCACA 906
Qy 901 CAACCAAGCACGCCACAGCAAGGAGAAACACACAAACCATTTCCCAAGATGCTGTGACT 960
Db 907 CAACCAAGCACGCCACAGCAAGGAGAAACACACAAACCATTTCCCAAGATGCTGTGACT 966
Qy 961 GAACTAGACAAATAACACAACTGCGACACCGTCCATGCCCCCTCATAACACTACCAACA 1020
Db 967 GAACTAGACAAATAACACAACTGCGACACCGTCCATGCCCCCTCATAACACTACCAACA 1026
Qy 1021 ATCTCTACTAACAACACCTCCAAACACAACTTCAGCACTCTCTGCGACCAATACAAAC 1080
Db 1027 ATCTCTACTAACAACACCTCCAAACACAACTTCAGCACTCTCTGCGACCAATACAAAC 1086
Qy 1081 ACCACCAATGACACACAGAGCAACATCACTGAAATGAGCAACACAGTGCCCGCTCG 1140
Db 1087 ACCACCAATGACACACAGAGCAACATCACTGAAATGAGCAACACAGTGCCCGCTCG 1146
Qy 1141 ATAACAACCTTGCCTCCAAACGGAATCCACACAGCAAGAGAGCAGCAGCAAAAAA 1200
Db 1147 ATAACAACCTTGCCTCCAAACGGAATCCACACAGCAAGAGAGCAGCAGCAAAAAA 1206
Qy 1201 GGGCCGCGCACAAACGCGACCAACAGCAAAATGAGCAATTTCAACAGTCTCCCGCCACC 1260
Db 1207 GGGCCGCGCACAAACGCGACCAACAGCAAAATGAGCAATTTCAACAGTCTCCCGCCACC 1266
Qy 1261 CCCAGTCCGACTGCAACACATCTGTATATTTTCAGAGAAAGCGA 1305
Db 1267 CCCAGTCCGACTGCAACACATCTGTATATTTTCAGAGAAAGCGA 1311
```

RESULT 10

ABT13454
ID ABT13454 standard; DNA; 2051 BP.

XX AC ABT13454;
XX

DT 30-JAN-2003 (first entry)

DE DNA encoding a chimeric filovirus protein MBGV-GP1/GP2.

XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.

OS Unidentified.
OS Chimeric.

XX PN W0200279239-A2.

XX PD 10-OCT-2002.

XX PF 31-JAN-2002; 2002WO-US003339.

XX 31-JAN-2001; 2001US-0267522P.
XX (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX Grogan CC, Hevey MC, Schmaljohn AL;
XX WPI; 2003-040651/03.
XX P-PSDB; ABJ18476.
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
XX useful for inducing an immune response against infection of different
XX filoviruses, specifically against both Ebola and Marburg viruses.
XX Disclosure; Page 81-84; 94pp; English.
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
XX comprising GP1 and GP2, where GP1 is from a filovirus different than that
XX of GP2. The chimeric filovirus GP protein is useful for inducing an
XX immune response against infection of different filoviruses, specifically
XX against both Ebola and Marburg viruses by being used as a vaccine. This
XX polynucleotide sequence represents a DNA encoding a chimeric filovirus
XX protein of the invention
XX Sequence 2051 BP; 655 A; 520 C; 400 G; 476 T; 0 U; 0 Other;

Query Match 69.2%; Score 1274; DB 10; Length 2051;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 12; Gaps 1;

Qy 1 ATGAAGACCCATGCTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
Db 10 ATGAAGACCCATGCTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 69
Qy 61 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGATTCGGTATGCTCCGGAAT 120
Db 70 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGATTCGGTATGCTCCGGAAT 129
Qy 121 CTCAGAGACAGAAAGCGTCCATCTGATGGGATTCACATGAGTGGGCAAAAATGTCGT 180
Db 130 CTCAGAGACAGAAAGCGTCCATCTGATGGGATTCACATGAGTGGGCAAAAATGTCGT 189
Qy 181 GATTTCCCTTTGGAGGATCCAGCGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db 190 GATTTCCCTTTGGAGGATCCAGCGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 249
Qy 241 GTTCAGTACACAGAGGGGAGAGCCAAACATGCTACAATAAAGTGTAAACCGATCCC 300
Db 250 GTTCAGTACACAGAGGGGAGAGCCAAACATGCTACAATAAAGTGTAAACCGATCCC 309
Qy 301 TCTCGAAAAATCCTTTGCTTTAGATCCTCTCAACCAATCCGCTGATCCGAAATGCAAA 360
Db 310 TCTCGAAAAATCCTTTGCTTTAGATCCTCTCAACCAATCCGCTGATCCGAAATGCAAA 369
Qy 361 ACTATCCATCATTTCAAGTCAAAACCTCATCGACAGGGGATCGCCCTTCATTATGG 420
Db 370 ACTATCCATCATTTCAAGTCAAAACCTCATCGACAGGGGATCGCCCTTCATTATGG 429
Qy 421 GGAGCATTTTTCTGTATGATCGCATTTGCCCTCCCAACAATGTACCGAGGCAAAATGCTTC 480
Db 430 GGAGCATTTTTCTGTATGATCGCATTT-----ATGTACCGAGGCAAAATGCTTC 477
Qy 481 ACTGAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCG 540
Db 478 ACTGAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCG 537
Qy 541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTAATAATATGACAGAT 600
Db 538 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTAATAATAATATGACAGAT 597
Qy 601 AGTAACGGAACGCAACGAATGACACTGGATGTTTTCGCGCTCTTCAAGAAATACAAATCT 660
Db 598 AGTAACGGAACGCAACGAATGACACTGGATGTTTTCGCGCTCTTCAAGAAATACAAATCT 657

```
QY 661 ACAAGAACCAACATGTGCTCGTCCAAAATACTTCCACCACTGCGCCACAGCCCGTCCG 720
D 658 ACAAGAACCAACATGTGCTCGTCCAAAATACCTTCCACCACTGCGCCACAGCCCGTCCG 717
QY 721 GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCAATACGAGCCCAAGC 780
D 718 GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCAATACGAGCCCAAGC 777
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAGAGGAAACCCACACA 840
D 778 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAGAGGAAACCCACACA 837
QY 841 ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA 900
D 838 ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA 897
QY 901 CAACCAAGACCGCCACAGACAGAGAGGAAACACACAAACCCATCCCAAGATCTGTGACT 960
D 898 CAACCAAGACCGCCACAGACAGAGGAGGAAACACACAAACCCATCCCAAGATCTGTGACT 957
QY 961 GAACTAGACAAAATAAACAACCTCCAAACACAACTTTCAGCACTCTCTGCGACCATTAACAAAC 1080
D 958 GAACTAGACAAAATAAACAACCTCCAAACACAACTTTCAGCACTCTCTGCGACCATTAACAAAC 1017
QY 1021 ATCTCTACTAACAACACCTCCAAACACAACTTTCAGCACTCTCTGCGACCATTAACAAAC 1080
D 1018 ATCTCTACTAACAACACCTCCAAACACAACTTTCAGCACTCTCTGCGACCATTAACAAAC 1077
QY 1081 ACCACCAATGACACACACAGAGACCAATCACTGAAATGAGCAACACAGTGCCTCCCTCG 1140
D 1078 ACCACCAATGACACACACAGAGACCAATCACTGAAATGAGCAACACAGTGCCTCCCTCG 1137
QY 1141 ATAAACAACCTGCTCCAAACGGGAATCCACACAGAGAGAGACACAGCAGCAGCAAAA 1200
D 1138 ATAAACAACCTGCTCCAAACGGGAATCCACACAGAGAGAGACACAGCAGCAGCAAAA 1197
QY 1201 GGGCCCGCCACAAACGGCACCACCAACAGCAAAATGAGCATTTTCCAGTCTCTCCCGCCACC 1260
D 1198 GGGCCCGCCACAAACGGCACCACCAACAGCAAAATGAGCATTTTCCAGTCTCTCCCGCCACC 1257
QY 1261 CCCAGTCTGACGACACAACTCTGTATATTTTCAGAGAAAGGATCG 1308
D 1258 CCCAGTCTGACGACACAACTCTGTATATTTTCAGAGAAAGGATCG 1305

RESULT 11
ADM48345
ID ADM48345 standard; DNA; 19112 BP.
AC ADM48345;
XX
XX
XX 18-NOV-2004 (first entry)
XX
XX Marburg virus viral protein genomic DNA.
DE Marburg virus
XX
XX Filovirus; viral protein; NP protein; VP30 protein; VP35 protein;
KW VP40 protein; L protein; membrane-associated protein; matrix protein;
KW polymerase complex protein; minor nucleoprotein; glycoprotein; GP;
KW viral mutagenesis study; vaccine; gene therapy; gene; ds.
XX
OS Marburg virus.
XX
XX Key
PH Location/Qualifiers
PT 103..2190
FT /*tag= a
FT /product= "NP protein"
FT 2944..3933
FT /*tag= b
FT /product= "VP35 protein"
FT 4567..5478
FT /*tag= c
FT /product= "VP40 protein"
```

```
FT CDS 5940..7985
FT /*tag= d
FT /product= "GP protein"
FT 8867..9712
FT /*tag= e
FT /product= "VP30 protein"
FT 10205..10966
FT /*tag= f
FT /product= "VP24 protein"
FT 11479..18474
FT /*tag= g
FT /product= "L protein"
XX
XX US2003215794-A1.
PN 20-NOV-2003.
XX
XX 29-JAN-2003; 2003US-00353856.
XX
XX 31-JAN-2002; 2002US-0353972P.
XX
XX (KAWA/) KAWAOKA Y.
PA (JASE/) JASENOSKY L D.
PA (NEUM/) NEUMANN G.
XX
XX Kawaoka Y, Jasenosky LD, Neumann G;
PI WPI; 2004-010776/01.
XX
XX P-PSDB; ADM48338, ADM48339, ADM48340, ADM48341, ADM48342, ADM48343,
XX ADM48344.
XX
XX Preparing filovirus useful in producing vaccines and gene therapy
XX vectors, comprises contacting a cell with a vector comprising a promoter
XX operably linked to a filovirus genomic cDNA or protein.
XX
XX Disclosure; SEQ ID NO 27; 104pp; English.
XX
XX The invention relates to a method of preparing filovirus. The method
XX involves contacting a cell with a vector comprising a promoter operably
XX linked to a filovirus genomic cDNA or its portion linked to a
XX transcription termination sequence or to a DNA segment encoding a
XX filovirus RNA transcriptase-polymerase, filovirus NP, VP30 or VP35, to
XX yield infectious filovirus. The portion of the cDNA when transcribed
XX yields a RNA capable of being packaged into filovirus virions or capable
XX of being replicated in the presence of filovirus proteins. The method is
XX useful for producing filovirus useful in viral mutagenesis studies and in
XX the production of vaccines and gene therapy vectors. The present sequence
XX is marburg virus (NP; L; polymerase complex-VP35; matrix protein-VP40;
XX glycoprotein-GP; minor nucleoprotein-VP30; membrane-associated protein-
XX VP24) genomic DNA.
XX
XX SQ Sequence 19112 BP; 6176 A; 3757 C; 3578 G; 5601 T; 0 U; 0 Other;
XX
XX Query Match 62.4%; Score 1148.2; DB 13; Length 19112;
XX Best Local Similarity 92.5%; Pred. No. 0;
XX Matches 1207; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 1 ATGAAGACCACATGTTCTCTTATCAGTCTTATTAATTAAGGGACAAAATCTCCCC 60
D 5940 ATGAAGACCACATGTTCTCTTATCAGTCTTATTAATTAAGGGATAAAATCTCCCT 5999
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGCGATTCGGTATGCTCCGAACT 120
D 6000 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGCGATTCGGTATGCTCCGAACT 6059
QY 121 CTCAGAGACAGAGACGTCCTCATCTGATGGATTCACACTGAGTGGGCAAAAATTCGT 180
D 6060 CTCAGAGACAGAGATGTCCTCATCTGATGGATTCACACTGAGTGGGCAAAAATTCGT 6119
QY 181 GATTCCCTTTGGAGGCATCCAGCGATGGGCTTTTCAGACAGAGTGTACCTCCCAAGAT 240
D 6120 GATTCCCTTTGGAGGCATCCAGCGATGGGCTTTTCAGACAGAGTGTACCTCCCAAGAT 6179
```

```
QY 241 GTTGTATACAGAGGGGAGGAGCCAAACATGCTACATATATAGTGTACGGATCCC 300
DB 6180 GTTGTATACAGAGGGGAGGAGCCAAACATGCTACATATATAGTGTACGGATCCC 6239
QY 301 TCTGGAATAATCTGCTGTGTAGATCCTCTCAACATCCGCTGACTATCCGAAATGCAAA 360
DB 6240 TCTGGAATAATCTGCTGTGTAGATCCTCTCAACATCCGCTGACTATCCGAAATGCAAA 6299
QY 361 ACTATCCATCATATTTCAAGGTCAAACCCCTCATGTCACAGGGGATCCGCTTCTCATTTATGG 420
DB 6300 ACTATCCATCATATTTCAAGGTCAAACCCCTCATGTCACAGGGGATCCGCTTCTCATTTATGG 6359
QY 421 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAATGTACCGAGGCAAGTCTTC 480
DB 6360 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAATGTACCGAGGCAAGTCTTC 6419
QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATAAGACAGTGCACAAAATGATTTCTTCG 540
DB 6420 ACTGAAGGGAACATAGCAGCTATGATTTGTCATAAGACAGTGCACAAAATGATTTCTTCG 6479
QY 541 CGGCAAGGACAAAGGTACCGTCATATGATCTGACTTCTTACTTAATAATATTTGGCAAGT 600
DB 6480 AGGCAAGGACAAAGGTACCGTCATATGATCTGACTTCTTACTTAATAATATTTGGCAAGT 6539
QY 601 AGTAAGGGAACGCAACGAAATGACACTGATGTTTCGGGGCTCTTCAAGAAATACAAATCT 660
DB 6540 AACTAAGGGAACGCAACGAAATGACACTGATGTTTCGGGGCTCTTCAAGAAATACAAATCT 6599
QY 661 ACAAGGAACCAACATGCTCCGTCCAAAATACCTCCACACTCCGACAGCCCGCTCG 720
DB 6600 ACGAAGAAATCAAAATGCTCCGTCCAAAATACCTCCACACTCCGACAGCCCGCTCG 6659
QY 721 GAGATCAAACTCAAGCAACCCCAACTGATGCCACCAAACTCAATPACACGAGCCCAAGC 780
DB 6660 GAGATCAAACTCAAGCAACCCCAACTGATGCCACCAAACTCAATPACACGAGCCCAAGC 6719
QY 781 AGTGTATGAGGAGCTTGCACATCTCGGCTCAGGGTCGGAGAAACGAGAAACCCACACA 840
DB 6720 AATGTATGAGGAGCTTGCACATCTCGGCTCAGGGTCGGAGAAACGAGAAACCCCTATACA 6779
QY 841 ACTTCTGATGGGTCACCAAGCAAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACC 900
DB 6780 ACTTCTGATGGGTCACCAAGCAAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACC 6839
QY 901 CAACCAAGCAGCCACAGCAAGGAGGAAACCAACAAACCACTTCCCAAGATGCTGTGACT 960
DB 6840 CAACCAAGCAGCCACAGCAAGGAGGAAACCAACAAACCACTTCCCAAGATGCTGTGACT 6899
QY 961 GAACTAGACAAAATAACAACTGACAACTGACAACTGACAACTGACAACTGACAACTGACAA 1020
DB 6900 GAACTAGACAAAATAACAACTGACAACTGACAACTGACAACTGACAACTGACAACTGACAA 6959
QY 1021 ATCTCTACTTAAACACACCTCCAAACACAACTTCAAGAAATGAGCAAAACCAAGTCCCTCG 1080
DB 6960 ATCTCTACTTAAACACACCTCCAAACACAACTTCAAGAAATGAGCAAAACCAAGTCCCTCG 7019
QY 1081 ACCCAATATGACAAACACAGAGCAAACTCACTGAAATATGAGCAAAACCAAGTCCCTCG 1140
DB 7020 ACCCAATATGACAAACACAGAGCAAACTCACTGAAATATGAGCAAAACCAAGTCCCTCG 7079
QY 1141 ATAAACACCTGCTCCAAACGGAATCCCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 7080 AAAACACCTGCTCCAAACGGAATCCCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 7139
QY 1201 GGGCCGCAACAGGCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB 7140 GGGCCGCAACAGGCAACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7199
QY 1261 CCAGCTGCTGCTGCAACATCTTGTATATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1305
DB 7200 CCAGCTGCTGCTGCAACATCTTGTATATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7244
```

```
RESULT 12
ABT13453
ID ABT13453 standard; DNA; 2046 BP.
XX AC ABT13453;
XX DT 30-JAN-2003 (first entry)
XX DNA encoding a chimeric filovirus protein RVN-GP1/MUS-GP2.
XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
XX immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX Unidentified.
XX OS Chimeric.
XX PN WO200279239-A2.
XX PD 10-OCT-2002.
XX PF 31-JAN-2002; 2002WO-US003339.
XX PR 31-JAN-2001; 2001US-0267522P.
XX (USSA ) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX Grogan CC, Hevey MC, Schmaljohn AL;
XX WPI; 2003-040651/03.
XX P-PSDB; ABJ18475.
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
XX useful for inducing an immune response against infection of different
XX filoviruses, specifically against both Ebola and Marburg viruses.
XX Claim 18; Page 78-80; 94pp; English.
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
XX comprising GP1 and GP2, where GP1 is from a filovirus different than that
XX of GP2. The chimeric filovirus GP protein is useful for inducing an
XX immune response against infection of different filoviruses, specifically
XX against both Ebola and Marburg viruses by being used as a vaccine. This
XX polynucleotide sequence represents a DNA encoding a chimeric filovirus
XX protein of the invention
SQ Sequence 2046 BP; 662 A; 486 C; 379 G; 519 T; 0 U; 0 Other;
Query Match 42.4%; Score 780; DB 10; Length 2046;
Best Local Similarity 74.8%; Pred. No. 5.2e-223;
Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;
QY 1 ATGAAGACCACATCTTCTTATCAGTCTTATCTTAAATTCAGGACACAAAAATCTCCCC 60
DB 1 ATGAAGACCACATATATTTCTGATTTAGTCTCATTTTATTCAGATATATAAACTCTCCCT 60
QY 61 ATTTTATGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
DB 61 GTTTTATGAATTTCTAGTAACAGCAACCTCAAGATGTAGATTCACTGTGCTCCGGAAC 120
QY 121 CTCAGAAAGACAGAAAGAGCTTCATCTGATGGATTTCACATGAGTGGCAAAAGTTGCT 180
DB 121 CTCGAAAGACAGAAAGATGTTTCTATGATGGATTTCACATGAGTGGCAAAAGTTGCT 180
QY 181 GATTTCCTTTTGGAGGATCCAGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 240
DB 181 GATTTCCTTTTGGAGGATCCAGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 240
QY 241 GTTGTATACAGAGGGGAGGAGCCAAACATGCTACATATATAGTGTACGGATCCC 300
DB 241 GTTGTATACAGAGGGGAGGAGCCAAACATGCTACATATATAGTGTACGGATCCC 300
QY 301 TCTGGAATAATCTTGTGTGTAGATCTCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360
```

```
Db 301 TCTGGAATACTTGCTGCTGGATCTCCAGTAATATATCCGCGATTACCCCTAAATGTAAA 360
Qy 361 ACTATCCATCATATCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTGTTTCATATATCAAGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Qy 421 GGAGCATTTTTCTGTATGATCGCATGCTCCCAACAACATGTACGAGGCAAAATCTTC 480
Db 421 GGGGCATTTTCTGTATGATCGCTGCTCTACAACAATGTACGAGGCAAGTCTTC 480
Qy 481 ACTGAGGGAATATAGCAGCTATGATTTGTAATAAGACAGATTCACAGAATGATTTTCT 540
Db 481 ACTGAGGGAATATAGCAGCTATGATTTGTAATAAGACAGATTCACAGAATGATTTTCT 540
Qy 541 CGGCAAGGACAGGGTACCGTCAATGATCTGACTTCTACTAATAAATATTTGGCAAGT 600
Db 541 AGGCAAGGACAGGGTATGCTCATGAATCTGACTTCTACTAATAAATATTTGGCAAGC 600
Qy 601 AGTAACGGAAACGAAACGAATGACACTGGATGTTTGGCGCTCTTCAAGAAATACAATTCT 660
Db 601 AGCAATGAACCGACAGAGAAATGATAGGGATGTTTGGCATCTCCAGAAATACAATCTCC 660
Qy 661 ACAAGAACCAACATGCTGCTCGTCCAAATACCTCCACACTGCGCCACAGCCCGTCGG 720
Db 661 ACAAAACATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG 720
Qy 721 GAGATCAAACTCACAAGACCCCAACTGATGCCAACCAACTCAATACACGACGCCAAGC 780
Db 721 AGCATTCCTCTACAAATCTCAATTAATATCTGTAATCTGGAACATATGAACCAAGT 780
Qy 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
Db 781 AGCGAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Qy 841 ACTTCTGATCGGTCACCAAGCAGGGCTTTTATCAAAATGCCACCCACTCCCTCACCAC 900
Db 841 ACTCTTAATGTATCTCACTGAACAGAAACAATCGTCAACAATATTTGTCACCTCTCTCACTA 900
Qy 901 CAACCAAGCAGCCACAGCAGGAGGAAACAACAACAACCAATTCCTCAAGATCTGTGACT 960
Db 901 CATCCAGACCTCTCAACATGAGCNAACAGTACGATCTTCCGACATGCTGTAACT 960
Qy 961 GAATCAGACAAATAAACAACATGTCACAAACCGTCCATGCCCTCATATAACTACCAACA 1020
Db 961 GAGCAATGGAACCGACCAACAACAACAGCAACGCTCTCAACAATACTAATAACA 1020
Qy 1021 ATCTCTACTAACACACCTCCAAACACACTTTCAGCACTCTCTGCAACCATTAACAAC 1080
Db 1021 ACTCCACCTATAACACTCTCAAGTACAACCTCAGTACTCTTCCCTCCCAACCCGCAAC 1080
Qy 1081 ACCACCAATGACACACAGACGACAAATCACTGAAATGACAAACAGTCCGCCCTCG 1140
Db 1081 ATCAACCAATATGATACACAGTGNACTAGCAGAAAGGCAACAAACCAATGCTCAGTTG 1140
Qy 1141 ATAAACACCTGCTCTCAACGGGAAATCCCAACACAGCAAGAGCACCAGCAGCAAAAAA 1200
Db 1141 AACACAACTCTAGATCCACAGAAATCCCAACACAGGACAGACACCAACAGCAAC 1200
Qy 1201 GGGCCCGCCACAAACGACCAACAGCAAAATAGCAGATTTACCGTCTCCCTCCCGCAC 1260
Db 1201 AACATCATATGACGACATCAGATATAACAAGCAAAACACCCCAACAATTTCTTCTCGGAT 1260
Qy 1261 CCAGCTCCAGTCGCAACATCTGTATATTTTCAAGAAAGCGATCG 1308
Db 1261 TCTAGTCCGACAAACCCGCCCTCTATATATCTTTAGAAAGAAACGATCG 1308
```

RESULT 13

ABT13456

ID ABT13456 standard; DNA; 2046 BP.

XX

AC

XX

```
DT 30-JAN-2003 (first entry)
XX DNA encoding a chimeric filovirus protein RVN-GP1/GP2.
XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
XX immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX Unidentified.
XX Chimeric.
XX WO200279239-A2.
XX 10-OCT-2002.
XX 31-JAN-2002; 2002WO-US003339.
XX 31-JAN-2001; 2001US-0267522P.
XX (USSA ) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX Grogan CC, Hevey MC, Schmaljohn AL;
XX WPI; 2003-040651/03.
XX P-PSDB; ABU18478.
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
XX useful for inducing an immune response against infection of different
XX filoviruses, specifically against both Ebola and Marburg viruses.
XX Disclosure; Page 89-91; 94pp; English.
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
XX comprising GP1 and GP2, where GP1 is from a filovirus different than that
XX of GP2. The chimeric filovirus GP protein is useful for inducing an
XX immune response against infection of different filoviruses, specifically
XX against both Ebola and Marburg viruses by being used as a vaccine. This
XX polynucleotide sequence represents a DNA encoding a chimeric filovirus
XX protein of the invention
XX Sequence 2046 BP; 668 A; 481 C; 383 G; 514 T; 0 U; 0 Other;
```

Query Match 42.4%; Score 780; DB 10; Length 2046;

Best Local Similarity 74.8%; Pred. No. 5.2e-223;

Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

```
Qy 1 ATGAGACACATGTTTCTTATCATGCTTATCTTAATTCAGGACAAAAATCTCCCT 60
Db 1 ATGAGACACATATTTCTGATGATCTCATTTTAAATCCAAAGTATAAAATCTCCCT 60
Qy 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGTCCTCGAAT 120
Db 61 GTTTTAGAAATGCTAGTAAGACCAACCTCAAGATGTAGATTAGTGTCTCCGNAAC 120
Qy 121 CTCGAGACAGAGAGAGCTCCTGATGGGATTCACATGAGTGGGCAAAAGTTGCT 180
Db 121 CTCGAGACAGAGAGAGTTCATCTGATGGGATTTACACTGAGTGGGCAAAAGTTGCT 180
Qy 181 GATTCCCTTTGGAGGCATCCAAGCATGGGCTTTTCAGACAGGTACTCTCCCAAGAT 240
Db 181 GATTCCCTTTGGAGGCATCCAAGCATGGGCTTTTCAGACAGGTACTCTCCCAAGAT 240
Qy 241 GTTAGTACACAGAGGGGAGAGCCAAACATGCTACAATATAAGTGTAAAGGATCC 300
Db 241 GTTAGTATACGAGAGGAGAGAGCCAAACATGCTACAATATAAGTGTAAAGGATCC 300
Qy 301 TCTGAAAAATCTTCTGTTAGATCTCTCTACCAACATCCGTCGACTATCCGAAATGCAAA 360
Db 301 TCTGAAAAATCTTCTGTTAGATCTCTCTCCAGTAATATCCGCGATTTACCTAAATGTAAA 360
Qy 361 ACTATCCATCATATTCAGGTCAAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTGTTTCATATCAAGGTCAAAACCTCATGCACAGGGGATTTGCCCTTCATTTGTTGG 420
```


Qy	421	GGAGCATTTTTTCTGTATGATCGCATTTGGCTCTCAACAACATGTATCCGAGGCAAAAGTCTTC	480
Db	421	GGGGCATTTTTTCTGTATGATCGCGTTGGCTCTCAACAATGTATCCGAGGCAAGTCTTC	480
Qy	481	ACTCAAGGGAAACATAGCAGCTATCATTTGTCTAATAAGACAGTGCACAAATCATTTTCTCG	540
Db	481	ACTGAGGAATAATAGCAGCTATGATTTGTTAATAAGACAGTTTCACGAATGATTTTTTCT	540
Qy	541	CGGCAAGGCAACAGGGTACCGTTCATATGAATCTGACTTCTCTAATAATAATTTGGACAAGT	600
Db	541	AGGCAAGGCAACAGGTTATCGTTCATGAACTTTGACCTCCACCANTAAATTTGGCAAGC	600
Qy	601	AGTAAACGGAAACGCAAAACGAATGACACTTGGATGTTTGGCGCTCTTCAAGAAATACAATTTCT	660
Db	601	AGCAATGAAACCGCAGAGAAATGATACGGGATGTTTTTGGCATCTCTCCAGAAATACAATCTCC	660
Qy	661	ACAAGAACCAAAACATGTCTCCGTCCAAATATACCTCCACACATGCCCCACAGACCGCTCCG	720
Db	661	ACAACAATCAAAACATGCCCTCCATCTCTTTAAACCTCCATCTCCCTGCCACAGTAACCTCG	720
Qy	721	GAGATCAAACTCACAGCACCCCAACTGATGCCACCAAACTCAATACCAACGAGCCCAAGC	780
Db	721	AGCATTCATCTCAAAATACCTCAAAATTAATCTGTTAATCTGGAACTATGAACCCCAAGT	780
Qy	781	AGTGATGATGAGGACCTCGCAACATCTCGGCTCAGGGTCGGGAGAACGAGAACCCCAACA	840
Db	781	AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAAACAGGGGCCCCACACA	840
Qy	841	ACTTCTGATGGGTTACCAAGCAAGGGTTTTCATCAACAAATGCGACCCACCTCCCTCACCA	900
Db	841	ACTCTTAATGTAGTCACTGTAAACAGAAAACAATCGTCAACAATATTGTCTCCACTCTTCACTA	900
Qy	901	CAACCAAGCGCCACACAGCAGGAGGAAACAACAACCACTATTCCTCAAGATGCTGTAGCT	960
Db	901	CATCCAAGCACCTCACAAACATGAGCAAAACAGTACGAATCTCTCCCGACATGCTGTAACT	960
Qy	961	GAACTAGACAAAAATAACAACCTGCACAAACCGTCCATGCGCCCTCTATAAACACTACCA	1020
Db	961	GAGCACAATGGAAACGACCCCAACAACAACGAGGACGCTCTCTCAACATACTAATAACA	1020
Qy	1021	ATCTCTACTAAACAACCTCCAAAACAACCTTTCAGCACTCTCTCTGCAACCAATTAACAAAC	1080
Db	1021	ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCTCCCTCCCAACCCGCAAC	1080
Qy	1081	ACGACCAATGACACACACAGACGACCAATCACTGAAATGAGCAAAACGAGTGCCCCCTCG	1140
Db	1081	ATCACCAATATATGATACAAACGTTGAACTAGCAGAAAGCGCAACCAACCAATGCTCAGTTG	1140
Qy	1141	ATACCAACCTCGCTCCAAACGGGAAATCCCAACCAACGACAGACAGAGGCACCAGCAGCAAAA	1200
Db	1141	AACACAACCTTAGATCCAAACAGAAATCCCAACCAACGACAGACCAACCAACGACCAACC	1200
Qy	1201	GGCCCCGCCCAACAGGGCACCAACAACGACAAATGAGCATTTTTCACAGTCTCTCCCCCACC	1260
Db	1201	AACATCATCATGACGACATCAGATATACACAGCAAAACACCCCAACAATTTCTTCTCCGGAT	1260
Qy	1261	CCCAGCTCGACTGCACAAACATCTTGTATATTTTCAGAGAAAGCGATCG	1308
Db	1261	TCTAGTCCGACAAACCGCGCTCTCTATATCTTTTAGAAGAAACGATCG	1308

RESULT 14

[illegible]

361	Qy	ACTATCCATCATATTC	AAAGTCAAAACCC	TCTATG	CCACAGGGAGTCG	CCCTTCATTTAT	TATGG	420
421	Qy	GGAGCAATTTTTCT	GTATGATCCAT	TGCTCCACAA	CATGTATCCGAGGACGA	AGTCTTCC	480	
481	Qy	ACTGAAGGGAA	CATAGCAGCTAT	GATTTGTCAA	TAAAGACAGTGCACAAAA	TGATTTTCTCG	540	
541	Qy	CGGCAAGGACA	AGGGTACCTCAT	TATGAATCTG	ACTTCTACTAAT	ATAATTTGGACAAGT	600	
601	Qy	AGTAA	CGGAACGAA	CGAATGAC	ACTGAGTGTTCGGCGCTCTT	CNAGGAATACAATCTT	660	
661	Qy	ACAAAGAAC	CAAAACATGTGCT	CCGTCCAAAAT	TACCTCCACAC	TGCCCCACAGCCGCTCCG	720	
721	Qy	GAGATCAA	CTCAAGACA	CCCCA	CTGTATGTCGCA	CAAACTCAATACCA	CGGACCCCAAGC	780
841	Qy	ACTTCTGAT	CGGTCACAA	GACAGGCGCTT	TCATCAACAAT	TGCCACCACTCCCTCACCA	900	
901	Qy	CAACCAAGCA	CGCCACAG	CAGGAGGAA	CAACAACAAC	CATTCCCAAGATGTGTGACT	960	
961	Qy	GAACTAGACA	AAATATACAA	CTGCAACA	CGCTCCATGCCCCCTCAT	ATACACTTACCACA	1020	
1021	Qy	ATCTCTACT	AAACAAC	ACTCTCAG	CACTCTCTCTG	CACCAATTACAAAAC	1080	
1081	Qy	ACCA	CAATGACA	ACACAGAC	CAATCATCTG	AAATGAGCAAA	CACTGTGCCCTCTG	1140
1141	Qy	ATA	CAACCCCTG	CTCCACG	GGGAATCCCA	CCACAGCAAGAC	CCACGACGCAAAA	1200
1201	Qy	GGCCCCGCCA	CAACGGCA	CCAAACAGCA	CAAA	TGAGCATTT	CACAGCTCTCCCCCCCACC	1260
1261	Qy	CCCAGCTCG	ACTGCA	CAACATCTT	GTATATTTCA	GAGAAAGCA	1305	
1305	Db	TC	TAGTCCGACAA	CCCCGCTCTCTAT	ACTTTTAGA	AAGAAACGA	1401	

RESUM.T 15

ACC71562

ACC71562
ID ACC71562 standard: DNA: 6902 BP.

AC ACC71562:

XX

DT 10-JUL-2003 (first entry)

XX

DE	VRC6703(pVR1012x/s Marburgdelta TM/h (codon optimised)) plasmid.
XX	
KW	Viruicide; vaccine; immune response; Ebola virus; Marburg virus;
KW	Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW	viral infection; filovirus; circular; cyclic; ds.
XX	
OS	Synthetic.
XX	
PN	WO2003028632-A2.
XX	
PD	10-APR-2003.
XX	
PF	24-SEP-2002; 2002WO-US030251.
XX	
PR	01-OCT-2001; 2001US-0326476P.
XX	
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX	
PI	WPI; 2003-371961/35.
DR	
XX	
PT	New bimodal priming and boosting compositions, useful as viral vaccines,
PT	specifically for eliciting an immune response against a filovirus or a
PT	disease caused by infection with filovirus.
XX	
PS	Claim 1; Page 213-215; 219pp; English.
XX	
CC	The present invention relates to a bimodal priming composition and
CC	boosting composition for priming and boosting an immune response to an
CC	antigen in an individual. The compositions comprise (a) a priming
CC	composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC	encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC	virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC	DNA plasmid selected from AC71521-AC71563 and (b) a boosting
CC	composition comprised of a replication-deficient adenovirus, comprising a
CC	nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC	paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC	epitope-bearing domain, or a replication deficient adenovirus selected
CC	from constructs AC71521-AC71563. The compositions are useful as viral
CC	vaccines, specifically for eliciting an immune response against a
CC	filovirus or a disease caused by infection with filovirus e.g. Ebola
XX	virus infection
XX	
SQ	Sequence 6902 BP; 1753 A; 1978 C; 1707 G; 1464 T; 0 U; 0 Other;
Query Match 37.4%; Score 688; DB 10; Length 6902;	
Best Local Similarity 70.5%; Pred. No. 4.5e-195;	
Matches 919; Conservative 0; Mismatches 385; Indels 0; Gaps 0	
Qy	1 ATGAGGACCATGTTTCCCTTATCAGTCTTATCTTAATTCAGGAGACAAAATCTCCCC 60
Db	1923 ATGAAGACCACTGCTGTTTCATCAGCTGATCTCTGATCCAGGGAATCAGACCTGCCC 198
Qy	61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTTCGGTATGCTCCGGAAT 120
Db	1983 ATCCTGAGATCGCAGCAACACACGCCCCAGAACGTGGACGCGTGTGCAGCGCAC 204
Qy	121 CTCAGAAAGACAGAAAGACGTCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
Db	2043 CTGCAGAAAGACCGAGGACGTGCACCTGATGGCTTCACCTTGAGCGGCGAGAAAGTGGCC 210
Qy	181 GATTCCCTTTGGAGGCATCCAAAGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db	2103 GACAGCCCTCTGGAGGCCGACGAAGAGTGGGCTTCAGGACCGGGCTGCCCTCCCAAGAAC 216
Qy	241 GTTCAGTACACAGAGGGGGAGGACCAAAACATGCTACAATATAAGTGTAAACGATCCC 300
Db	2163 GTGGAGTACACCGAGGGGAGGAGCCAGACCTGCTACAACATCAGCGTGNACCGACCCC 222
Qy	301 TCTGGAAAAATCCTTGCTGTAGATCCTCCTTAACCAACATCCGTGATATCCGAAATGCATA 360
Db	2223 AGCGGCAAGAGCCCTGCTGAGACCTTCCCAACCAACATCAGGACTACCTAAGTGCAAG 228

```
QY 361 ACTATCCATCATATTCAAGGTCAAAACCCCTCATGCAAGGGGATCGCCCTTCATTATTATGG 420
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2283 ACCATTCACCATCCAGGGCCAGAAACCTCTACGCCCAAGGCATCGCCCTGCACCTGTGG 2342
QY 421 GGAGCATTTTCTCTGTATGATCGCATTCCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2343 GCGGCTTCTCTGTATGACAGAGATCGCAGCACCAACATGTACAGGGGCAAGGTGTTTC 2402
QY 481 ACTAAGGGAAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTTTCTCG 540
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2403 ACCGAGGGCAACATCGCGCCATGATCGTTAAACAAGACCGTGCACAAGATGATCTTCAGC 2462
QY 541 CGGCAAGGACAGGGCTACCGTCATGATCTGACTTCTACTAATAAATATTGGACAAGT 600
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2463 AGGCAGGGCCAGGGCTACAGGCACATGAACCTGACAGCACCAACAGTACTTGGACCAGC 2522
QY 601 AGTAAACGGAACGCAAAACGAATGACACTGTGATGTTTCGGCGCTCTTCAAGAAATACAAATTCT 660
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2523 AACAAAGGCACCCAGACCAACGACACCGGCTGCTTCGGGGCCCTGCAGGAGTACAAACAGC 2582
QY 661 ACAAGAACCAGAAATGTGCTCCGTCCAAAATACCTCCACCACTGCCCCACAGCCCGTCCG 720
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2583 ACCAAGAACCAGACTCGCGCCCCAGCAAGATCCCCAGCCCCCTGCCACGCGCCAGGCC 2642
QY 721 GAGATCAAACTCACAGCAACCCCACTGATGSCCACCACCAACTCAATACCAAGGACCCAAAGC 780
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2643 GAGATCAAGCCACAGCACCCCCACCGACGCCCCCTGAAACACACCGACGCCCAAC 2702
QY 781 AGTGATGATGAGGAGCTCGCAACATCCGCTCAGGGTCCGGAGAACGAGAACCCACACACA 840
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2703 AACGACGACGAGACCTGATCACAGCGGCGAGCGGCGGAGCAGGAGCCCTTACACC 2762
QY 841 ACTTCTGATGCGGTCCACCAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA 900
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2763 ACCAGCGACGCGTGACCAAGCAGGGGCTTGAGCAGCACCATGCTCTTACCCCTAGCCCT 2822
QY 901 CAAACAGGACGCCACAGAGGAGGAAACACACAAACCATTTCCCAAGATGCTGTGACT 960
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2823 CAGCCGACGACCCCTCAGCAGGAGGGCAACACACCGACCAAGCCAGCGGCAACCGTGACC 2882
QY 961 GAACTAGACAAAATAACACAACTGCACAAACCGTCCATGCCCTCATAACTACACACACA 1020
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2883 GAGCCCAACAGACCAACACACCCGCCAGCCCGATGCTCTCTACACACACACCCGCC 2942
QY 1021 ATCTCTACTAACACACCTCCAAACACAACTTTCAGCACTCTCTGTGCACCATTTACAAAC 1080
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2943 ATCAGCACCAACACACACAGCAAGAACTTTCAGCACCCCTGAGCGTGCAGAAC 3002
QY 1081 ACCACCAATGACACACACAGAGCAATCACTGAAAATGAGCAAAACAGTGCCCCCTCG 1140
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3003 ACCACCAACTACGACACCCAGAGCACCGCCACCCGAGAACGAGCAGACAGCGCCCTAGC 3062
QY 1141 ATAACAACCTCGCTCCACAGGGGAATCCACCACAGCAAGAGCACCCAGCAGCAABAAA 1200
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3063 AAGACACCTCTGCTCCACCGGCAACCTGTGACACCGCAAGAGCACCAACACCAAC 3122
QY 1201 GGCCCCGCCACACAGGACCAACACGACAAATGAGCATTTTCAACAGTCTCTCCCCCACC 1260
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3123 GGCCCCACACACCGGCCCTTAACATGACCAAGGCCACCTGACCGCCCCAGCCCCCACC 3182
QY 1261 CCCAGCTGACTGCAACAATCTTGTATATTTTCAAGAAAGCG 1304
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3183 CCCAACCCCCACCCAGCACCTGTGTACTTTCAGGAAGAGAG 3226
```

Search completed: July 31, 2005, 16:57:16
Job time : 968.473 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:13:05 ; Search time 6009.96 Seconds
(without alignments)
11660.032 Million cell updates/sec

Title: US-10-066-506A-3

Perfect score: 1841

Sequence: 1 atgaagaccacatgttctc.....tttgtcttttagtgaattc 1841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hrc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	99	5.4	895	9	CNS0071A
C 2	90.4	4.9	1101	9	CNS00PXE
C 3	87.2	4.7	922	9	CNS0073W
C 4	83.4	4.5	884	9	CNS006U0
C 5	80.8	4.4	1225	9	CNS0166K
C 6	79.2	4.3	909	9	CNS000TL
C 7	75.6	4.1	1101	9	CNS000JT
C 8	75	4.0	1965	9	CL090566
C 9	73.6	3.9	939	9	CNS006U0
C 10	72.6	3.9	1101	9	CNS0172T
C 11	72.2	3.9	1787	9	CG754239
C 12	72.2	3.9	1626	9	AG131986
C 13	71.8	3.9	1455	9	AG382036
C 14	71.4	3.9	1101	9	CNS00LO0
C 15	70.2	3.8	1101	9	CNS017YH
C 16	70	3.8	1101	9	CNS006QP
C 17	69.2	3.8	902	9	CNS00075A
C 18	68.2	3.7	861	9	CNS000CNG
C 19	68	3.7	939	9	CNS000CNG
C 20	67.4	3.7	1559	9	CG7556460
C 21	67	3.6	796	8	B12592
C 22	67	3.6	993	9	CNS005N6
C 23	67	3.6	1319	9	AG371230
C 24	67	3.6	1667	9	CG7556610

C 25	66.8	3.6	913	6	CA986290
C 26	66.2	3.6	470	4	BG786319
C 27	66	3.6	723	8	AZ627893
C 28	66	3.6	1025	9	CL487193
C 29	65.6	3.6	527	9	CE480293
C 30	65.4	3.6	925	7	CK425533
C 31	65	3.5	769	8	BZ579780
C 32	65	3.5	869	9	AG137111
C 33	65	3.5	1147	8	CC187235
C 34	64.8	3.5	513	9	CNS02CMF
C 35	64.8	3.5	633	7	CV070818
C 36	64.8	3.5	860	9	CNS018FL
C 37	64.6	3.5	937	9	CNS006ST
C 38	64.6	3.5	1153	9	CNS07BVP
C 39	64.6	3.5	1345	8	AQ743328
C 40	64.6	3.5	1613	8	BZ557155
C 41	64.4	3.5	917	9	CL466449
C 42	64.2	3.5	1131	9	CNS03AWV
C 43	64	3.5	1350	9	AG393392
C 44	64	3.5	2263	3	CR698468
C 45	63.8	3.5	732	9	AG135778

ALIGNMENTS

RESULT 1
CNS0071A/c
LOCUS
DEFINITION
CNS0071A 895 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL066286

VERSION
GSS.

KEYWORDS
SOURCE

ORGANISM
Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 895)

AUTHORS
Genoscope.

TITLE
Direct Submission

JOURNAL
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

COMMENT
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Location/Qualifiers

1. 895

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR14B09"

/note="end : TET3"

ORIGIN

Query Match 5.4%; Score 99; DB 9; Length 895;

Best Local Similarity 24.1%; Pred. No. 6.1e-17;

[illegible]

RESULT 2	CNS00FXE	1101 bp	DNA	linear	GSS 03-JUN-1999
LOCUS	Drosophila melanogaster genome survey sequence TBT3 end of BAC:				
DEFINITION	BACR32C19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL071370				
VERSION	AL071370.1	GI:4951210			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
	Ephyarodea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				
AUTHORS	Genoscope.				
TITLE	Direct Submmission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequences :				
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
	- web : www.genoscope.cns.fr)				
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osogawa and Aaron Mammot in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial, EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .				
FEATURES	Location/Qualifiers				
	1..1101				
source	/organism="Drosophila melanogaster"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:7227"				
	/clone="BACR32C19"				
	/clone_lib="RPCI-98"				

ORIGIN /note="end : TET3"

Query Match 4.9%; Score 90.4; DB 9; Length 1101;
Best Local Similarity 27.5%; Pred.No. 2.1e-14;
Matches 108; Conservative 135; Mismatches 149; Indels 1; Gaps 1;

QY 872 CATCAACAATGCCACCCACTCTCCTCACCAACAAGCAGCGGCACGACGAAGGAGAAACA 931
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1096 MATAAMAHAYACAAAAAAMCMMAAACAAMMACMMAAMCMCMMAAMMM 1037

QY 932 ACACAAACCATTCCCAGATGCTGTGACTGTAAGTACGACAAAATAACACAACATGCACAAC 991
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1036 ACAGMH-MCKMKMMCAAHANMANAMAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 978

QY 992 CGTCATGCCCCCTCATACACTACCAATCTCTACTAACAAACACTCTCCAACACACAAC 1051
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 977 MAMMMNMCACAMCAAMAAAMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 918

QY 1052 TCAGCACTCTCTCTGCAACCAATTACAAACACCAACCAATGACAAACACACAGAGACACATCA 1111
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 917 CATAAMMTMMCMCAACMMAMMAAAAAAAAAAMMACMAAMCAMCAAMMAAAAAAAAAA 858

QY 1112 CTGAAATGAGAAACGAGTGCCCTCTGATTAACAACCTGCTCCACGGGAAATCCCA 1171
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 857 MMMAANMAACMAAMMMCMCMMAACMCMCMCMCMMAAMMMCMCMCMCMCMCMCMCMCMCM 798

QY 1172 CCACAGAAGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1231
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 797 MMMAACMMCMCSAGAMCMCMCAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMA 738

QY 1232 ATGAGCATTTTCACGAGTCTCTCCGCCCAACCCCA 1264
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 737 AMACCCCMGACCCCTCMTCCMCMCAACGCMMA 705

RESULT 3 CNS0073w/c 922 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL066784
VERSION AL066784.1 GI:4945247
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Rukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 922)

REFERENCE Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequences :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES Location/Qualifiers
source 1..922

```
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14N21"
/clone_lib="RPCI-98"
/note="end : TET3"
```

ORIGIN

```
Query Match      4.7%; Score 87.2; DB 9; Length 922;
Best Local Similarity 22.1%; Pred. No. 1.7e-13;
Matches 78; Conservative 145; Mismatches 130; Indels 0; Gaps 0;

Qy 930 CAACACAAACCATCCCAAGATGCTGCTGACTGACAAATACACAAATACACAAATGCACA 989
Db 905 MVMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 846

Qy 990 ACCGTCCATGCCCCCTATAACACTTACCAATCTCTACTAACACACCTCTCAACACAA 1049
Db 845 MVMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 786

Qy 1050 CTTGAGCACTCTCTCTGACCAATTAACAAACACCAACCAATGACACACAGAGACAAT 1109
Db 785 MVMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 726

Qy 1110 CACTGAAATGAGCAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1169
Db 725 MVMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 666

Qy 1170 CACACAGCAAGAGCAGCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1229
Db 665 AMAMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 606

Qy 1230 AATGAGCATTTTACCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1282
Db 605 MVMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 553
```

RESULT 4

```
CNS006U0/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14N21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
```

AL065923

AL065923.1 GI:4944891

GSS.

Drosophila melanogaster

Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 884)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mammosser in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain y2; cn bw sp, the same strain used for the BDGP's

p1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Location/Qualifiers

FEATURES

FEATURES

source

1..1225

/organism="Drosophila melanogaster"

1..884

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR14N21"

/clone_lib="RPCI-98"

/note="end : T7"

ORIGIN

```
Query Match      4.5%; Score 83.4; DB 9; Length 884;
Best Local Similarity 20.1%; Pred. No. 2.1e-12;
Matches 78; Conservative 171; Mismatches 138; Indels 1; Gaps 1;

Qy 855 CACCAAGCAGGCGTTTCATCAACAATGCCACCACTCCCTCACCACAAACCAACGACGCC 914
Db 881 MVMVMHHTTKKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKX 822

Qy 915 ACAGCAGGAGGAGAAACAACAACCAATTCCTCCAGATGCTGTGACTGAATCTAGACAAA 974
Db 821 MVMVMHHTTKKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKX 762

Qy 975 TAACACAACTGCACAAACCGTCCATGCCCCCTCATAAACACTACCAACAATCTCTACTAAC 1034
Db 761 MVMVMHHTTKKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKX 702

Qy 1035 CACCTCCAAACACAACTTCAGCACTCTCTGTGACCACTTACAAACACCAACCAATGACA 1094
Db 701 MVMVMHHTTKKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKXKX 643

Qy 1095 CACACAGCAGCACAATCACTGAAATGAGCAAAACAGTCCCTCGATTAACAACCTCTGCC 1154
Db 642 CCCCCCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 583

Qy 1155 TCACACGGGAAATCCACACAGCAAGAGACAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1214
Db 582 CAMCTCACCAAMCAAHMMAAAMAMCCABCMCMCHMASMCMCMACAYMCCCCCAMBTB 523

Qy 1215 GGCACCAACACAGCAAAATGAGCAATTC 1242
Db 522 KCBMCMYBCCCCCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 495
```

RESULT 5

CNS0166K/c

LOCUS

DEFINITION

Drosophila melanogaster

Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1225)

Genoscope.

Direct Submission

Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end

collaboration with the European Drosophila Genome Project (EDGP) -

<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC

library (Dros BAC) was made by Alain Billaud at CEPH (Centre

d'Etude du Polymorphisme Humain) with funding provided by a MRC

project grant. The DNA was prepared from embryos by Alain Bucheton

and Genevieve Payan. It has been constructed in the vector

pBeloBAC11.

Location/Qualifiers

Genoscope.
Direct Submission
Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segr@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammeter in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RP1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
 Location/Qualifiers
 1. .1101
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="BACR48P19"
 /clone_lib="RP1-98"
 /notes="end : TET3"

ORIGIN
Query Match 4.1%; Score 75.6; DB 9; Length 1101;
Best Local Similarity 13.8%; Pred. No. 4.3e-10;
Matches 60; Conservative 208; Mismatches 166; Indels 0; Gaps 0;
QY 868 CTTTTCATCAACATGCCACCCTCCCTCACCAACCAAGACGCGCCACAGAGGAGGA 927
DB 1096 HHTTMMMMCMCHTAAAMMMMMMMMMMMMAAMMMCMCMCMCMCMCMCMCMCM 1037
QY 928 AACACACAAACCAATTCCTCAAGATGCTGTGACTGAACATAGACAAAATAACACACTGCA 987
DB 1036 TMMMMMMMMMAAMMMMAAMMMMMMAATTTTTHMMMAAMMMHMMMMMAATTHAHTT 977
QY 988 CAACCGTCCATGCCCTCATACACTACACAACTCTCTACTACACACCTCCAAACAC 1047
DB 976 TTTTHTTMMAMCTTTTMMMMMAAMMAAMMMCMCMCMCMCMCMCMCMCMCMCMCM 917
QY 1048 AACTTCAGACTCTCTCGCACCATTAACAAAACACCACCAATGACACACACAGAGCACA 1107
DB 916 MMTTTHHHMMCMCMCMCCCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMMAA 857
QY 1108 ATCACTGAAATGAGCAAAACGATGCCCTCGATATAACAACCCCTCCCAACGGGAAAT 1167
DB 856 MMAAMMMMAAMMMMAAMMTTTTTTTHMMMAAMMMMAATTTMMMMMAAMMMMAATTTM 797
QY 1168 CCCACCAAGCAAGAGCACCAGCAGCAAAAAAGGCCGCCCAACACGGCACCACACG 1227
DB 796 MMAAMMMMAAMMMMAAMMMMAAMMMMAAMMMMAAMMMMAAMMMMAAMMMMAAMMM 737
QY 1228 ACAAATGAGCATTCACAGCTCTCCCTCCCAACCCCGAGCTGAGTGCAGCAACATCTTGA 1287
DB 736 TTMHMMMAAMMMTCHMMMTTHMMMTTTTMMMAAMMAAMMMMAAMMMMAAMMMMAA 677
QY 1288 TATTTCAGAGAA 1301
DB 676 TAATAAAAAAAA 663

RESULT 8
AG136151
LOCUS AG136151
DEFINITION Pan troglodytes DNA, clone: PTB-149K08.F, genomic survey sequence.
ACCESSION AG136151
VERSION AG136151.1 GI:16665829

```

GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM    Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE   1
AUTHORS     Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE       BAC end sequences of Library PTB
REFERENCE   2 (bases 1 to 837)
AUTHORS     Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
            Totoki, Y., Watanabe, H. and Sakaki, Y.
DIRECT SUBMISSION
TITLE       Direct Submission
JOURNAL     Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail: chimpansegsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,
            Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT     Clones are derived from the chimpanzee BAC library PTB. This BAC end
            was generated during the RED process and may have higher chance of
            clone tracking errors.
PRIMERS
SEQUENCING: -21M13
LIBRARY
Vector      : pKS145
R.Site 1    : SacI
R.Site 2    : SacI.
Location/Qualifiers
1. .837
   /organism="Pan troglodytes"
   /mol_type="genomic DNA"
   /db_xref="taxon:9598"
   /clone="PTB-149K08.F"
   /sex="male"
   /cell_type="lymphoblast"
   /clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES             source
source
Query Match          4.1%; Score 75; DB 9; Length 837;
Best Local Similarity 47.1%; Pred. No. 5.9e-10;
Matches 264; Conservative 0; Mismatches 292; Indels 5; Gaps 1;

Qy 727 AAATCTCACAAGCACCCCAACTGATGCCACCAAACTCAATACCAACGACCAACCAAGCATGTAT 786
    |||||
Db 261 AAACCACCAACCNACNCAACCAACAAACCAACCACTTATCCACCAACCCACCCACCAACCCACA 320
    |||||

Qy 787 GATGAGGACCTCGCAACATCGGCTCAGGTCGGGAGAGAGAGAACCCACACAACTTCT 846
    |||||
Db 321 AACACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 380
    |||||

Qy 847 GATGCGGTCCACCAAGCAAGGGCTTTTCATCAACAATGCCACCCCACTCCCTCACCACCAACCA 906
    |||||
Db 381 AACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 440
    |||||

Qy 907 AGCAGCGCCACAGCAAGGAGGAGAAACACACAAAACCAATTCGCCAGATGCTGTGACTGAACCTA 966
    |||||
Db 441 AGCAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 500
    |||||

Qy 967 GA-----CAAAAATACACAATGCGCAACCGGTTCATGCCCCCTCTATACACTTACACAA 1021
    |||||
Db 501 CAACCAACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 560
    |||||

Qy 1022 TCTCTACTAACACCTCCCAACACAACTTCAGCACTCTCTCTGCGACCACTTACAAAACA 1081
    |||||
Db 561 CCACAAAACCCACACATCTCCCAAAAACCAACCAACCAACCAACCAACCAACCAACCAAC 620
    |||||

Qy 1082 CCACCAATGACAAACACACAGAGCAACAATCACTGAAATAGACAAACCAAGTGCCTCCCTCGA 1141
    |||||
Db 621 CCAACACATCAACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 680
    |||||

Qy 1142 TAAACAACCTTGCTTCCAACGGGAAATCCCAACCAAGCAAGAGACGACGAGCAAAAAG 1201
    |||||

```

```

Db      681  ANCAACTCAACCAACCAACCAAAATCCAAAACAAACAAACAAACCAACCA 740
Qy      1202  GCCCGCCACAAACGGCACCACAAACAGACAAATGAGCATTTTACCAAGTCCTCCCGCCACCC 1261
Db      741  ACCAACCCCGCCCAACACACACCAACCAACCAACCAACCAACCAACCAACCAACCA 800
Qy      1262  CGAGTCGACTGCACAAATC 1282
Db      801  AAACCAAAACAAACAAACCAACC 821

RESULT 9
LOCUS   CL090566/c
DEFINITION
  ISBI-17017 T7.1 ISBI Xenopus tropicalis genomic clone ISBI-17017,
  genomic survey sequence.
ACCESSION
  CL090566
VERSION
  CL090566.1 GI:40584201
KEYWORDS
  GSS.
SOURCE
  Xenopus tropicalis (western clawed frog)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
  Xenopodinae; Xenopus; Silurana.
REFERENCE
  1 (bases 1 to 1965)
AUTHORS
  Krenitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
  Madris,E. and Wilson,R.
TITLE
  A physical map of the xenopus tropicalis genome
JOURNAL
  Unpublished (2003)
COMMENT
  Contact: Richard K Wilson
  Genome Sequencing Center
  Washington University School of Medicine
  Email: submissions@watson.wustl.edu
  Insert Length: 75000 Std Error: 0.00
  Seq primer: T7 TAATAGATCTACTATAGG
  Class: BAC ends
  High quality sequence start: 781
  High quality sequence stop: 832.

FEATURES
  source
    1..1965
    /organism="Xenopus tropicalis"
    /mol_type="genomic DNA"
    /db_xref="taxon:8364"
    /clone="ISBI-17017"
    /clone_lib="ISBI"
    /note="Vector: pBelobAC11; ISBI-1 Xenopus tropicalis BAC
    Library Segment 1"

ORIGIN
  Query Match 4.0%; Score 73.6; DB 9; Length 1965;
  Best Local Similarity 48.8%; Pred. No. 2e-09;
  Matches 199; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

Qy      875  CAACAATGCCACCCACTCCTCACCACAAACCAAGCAGCCACGAGGAGGAAACACA 934
Db      1185  CACCAACCAACCAACCAACCCCAACCAACCAACCAACCAACCAACCAACCAACCC 1126
Qy      935  CAACCAATTCACCAAGATGCTGTGACTAGACAGAAATACAACTGCACACCGT 994
Db      1125  ACACCACCAACCAACCCCAACCCCAACCAACCAACCAACCAACCAACCAACCAAC 1066
Qy      995  CCATGCCCCCTCATACACTACCACTCTCTACTAAACAAACCTTCCAAACACAACTTCA 1054
Db      1065  CCCCCACACACCCACACCCACACCAACAAAGCCACCCACCCACCAACCAACCAACC 1006
Qy      1055  GCATCTCTCTGGACCATTTACAAAACACCAACCAATGACAAACACACAGAGCACAATCTG 1114
Db      1005  CCCCCAAACAAACCCCAACCCCAACCAACCAACCAACCAACCAACCAACCAACCCCA 946
Qy      1115  AAAATGAGCAACCAAGTGCCTCGATACCAACCTTCCTCAACGGGAATCCACCA 1174
Db      945  CACACCAACAAACCAACCAACCAACCAACCAACCAACCCCAACCCCAACCAACCAACCC 886

```

```

Qy      1175  CAGCAAGAGGAGCAGCAGCAGCAAAAAGGCCCGCCACAAACGGCACCACCAACGACAAATG 1234
Db      885  CAACACGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 826
Qy      1235  AGCATTTTCACAGTCTCTCCCGCCACCCCGCCAGCTGCAGTGCACAAATC 1282
Db      825  AAAAAAAACACACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 778

RESULT 10
LOCUS   CNS006UU/c
DEFINITION
  Drosophila melanogaster genome survey sequence T7 end of BAC #
  BACR14L09 of RPCI-98 library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
ACCESSION
  AL065953
VERSION
  AL065953.1 GI:4944921
KEYWORDS
  GSS.
SOURCE
  Drosophila melanogaster (fruit fly)
ORGANISM
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 939)
AUTHORS
  Direct Submission
TITLE
  Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :
  BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
COMMENT
  Determination of this BAC-end sequence was carried out as part of a
  collaboration with the Berkeley Drosophila Genome Project (BDGP).
  The BDGP is constructing a physical map of the Drosophila
  melanogaster genome using these BACs. For further information
  please see http://www.fruitfly.org The BDGP Drosophila
  melanogaster BAC library was prepared by Kazutoyo Oseegawa and
  Aaron Mammoser in Pieter de Jong's laboratory in the Department of
  Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
  NY. The library is named RPCI-98 and was constructed by partial
  EcoRI digestion of Drosophila DNA provided by the BDGP from the
  isogenic strain Y2; cn bw ap, the same strain used for the BDGP's
  P1 and EST libraries. A more detailed description of the library
  and how to order individual BAC clones, the entire library, or
  filters for hybridization from the BACPAC Resource Center can be
  found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
  source
    1..939
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /db_xref="taxon:7227"
    /clone="BACR14L09"
    /clone_lib="RPCI-98"
    /note="end : T7"

ORIGIN
  Query Match 3.9%; Score 72.6; DB 9; Length 939;
  Best Local Similarity 33.8%; Pred. No. 3.1e-09;
  Matches 124; Conservative 92; Mismatches 150; Indels 1; Gaps 1;

Qy      795  CCTCGCAATCTCCGCTCAGGGTCGGAGACGAGACCCCAACCACTTCTGTATCGGT 854
Db      865  MSSCMSACCAACRCMGMAAGCGCCCGCCAGACGAGACGCGMGASVCGGASVGVCA 806
Qy      855  CACCAAGCAAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACCAACCAAGCAGCC 914
Db      805  VGSVRMACGAGSAGVSGVRCACACGVMVVCASCSMACAMMARMCAMAVASAAANAA 746
Qy      915  ACAGCAGGAGGAACAAACAAACCAATTCCTCCAGATGCTGTGACTGACTAGACAAA 974
Db      745  VAHVAGAVARGGGRAAMGRCGCAACRMAACMWCACAMAMARMSACGSAACAMMAACAS 686
Qy      975  TAAACAACCTGCACACCGTCCATGCCCGCTTCATAAACAACCTACCACTACTTACTA 1034
Db      685  VARMACAMAMACAMMAAAACSMRCVMCVSRACASMACAACCAACCAACRACACRCC 626

```

	QY	1112	CTGAAATGAGCAACACGAGTGCCTCGATATACAACCCCTGCTCCAACGGGAATGCCA	117
	Db	860	CCACACMACWACCAACAAAACAAAAAMCWAHMAACCAHAACCCCAACMAAACMMACC	801
	QY	1172	CCACAGAAGAGCACGACGAGCAAAAAGAGCCGCCACACACGCAACCAACACACGACAA	1231
	Db	800	CCCMCAMACMCYAMCAWCHVCTTTTTTTTWTACWYCWAAAHMAYAVATTACY	741
	QY	1232	ATGAGCATTTTCACAGTGCTCTCCCCCAGCCAGCTCGACTGCAACATCTTGTATATT	1291
	Db	740	CAAAAWAYYYYYYYMYCYCTCTHMATACVMCCYAHYVWTVYTWMTWTHAWAMTW	681
	QY	1292	TCAGAAGAGCGATCGGCAATTTGCTCAATCTCAACCCAAATCAACCCCTTAATTACATT	1351
	Db	680	YTATMCMCAAMCTMMCAAWTCATCCCTCTTCYCYYYTTACMYACCTKCCY	621
	QY	1352	ACT 1354	
	Db	620	YCT 618	
	RESULT 12			
	CG754239/c			
	LOCUS			
	DEFINITION	CG754239	1787 bp DNA linear GSS 24-OCT-2003	
		P049-3-D12.ya Ppa EcoRI BAC Library Pristionchus pacificus genomic,		
		genomic survey sequence.		
	ACCESSION	CG754239		
	VERSION	CG754239.1	GI:37979530	
	KEYWORDS	GSS.		
	SOURCE	Pristionchus pacificus		
	ORGANISM	Pristionchus pacificus		
		Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;		
	REFERENCE	Neodiplogasteridae; Pristionchus.		
	AUTHORS	1 (bases 1 to 1787)		
	TITLE	Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,		
		Buntjer,J., van der Weulen,M. and Sommer,R.J.		
		An integrated physical and genetic map of the nematode Pristionchus		
	JOURNAL	pacificus		
	MEDLINE	Mol. Genet. Genomics 269 (5), 715-722 (2003)		
	PUBMED	22835951		
	COMMENT	12884007		
		Contact: Sommer RJ		
		Evolutionary Biology		
		Max-Planck-Institute for Developmental Biology		
		Spemannstr. 37-39, Tuebingen D-72076, Germany		
		Tel.: 00497071601371		
		Fax: 00497071601498		
		Email: ralf.sommer@tuebingen.mpg.de		
		Class: BAC ends.		
	FEATURES	Location/Qualifiers		
	source	1..1787		
		/organism="Pristionchus pacificus"		
		/mol_type="genomic DNA"		
		/strain="California"		
		/db_xref="taxon:54126"		
		/clone_lib="Ppa EcoRI BAC Library"		
		/note="The library was generated by a partial digest of		
		the genomic DNA with EcoRI and cloning into the BAC		
		vector."		
	ORIGIN			
	Query Match	3.9%;	Score 72.2; DB 9; Length 1787;	
	Best Local Similarity	48.9%;	Pred. No. 5e-09;	
	Matches 194; Conservative	0; Mismatches 203; Indels 0; Gaps		
	QY	883	CCACCCACTCCCTCACCAACCAAGCAGCCACAGCAGAGGAGGAAACCAACAAACCAT	942
	Db	758	CCACCAACCATATCACACCAACACACACCCCTCCCATCAACACACACACACACCA	699
	QY	943	TCCCAAGATGCTGTGACTGAACCTAGACAAAATAACAACTGCACACCGTCCATGCC	1002
	Db	698	ACACCAAGAGACGACCTTAACACACAAACACCCCACCAAAAAAAAAAAGCCACCCC	639

```

QY 1003 CCTCATAACACTTACCAATCTCTACTACACACCTCCAAACAACTTTCAGCACTCTC 1062
Db 638 AACACCCCCCACAACACCAACCAAAACAAACCAACCAACCACTACACAAACACAC 579
QY 1063 TCTGCACCAATTACAAAACACCAACCAATGACACACAGAGACACAACTCAGTGAATGAG 1122
Db 578 TCAACACATCACAACATACACACACACCCCATACACTTCCACCCACACCCCCAC 519
QY 1123 CAACACAGTGGCCCCCTCGATACAAACCTGCTCCAAACGGGAAATCCCAACACAGCAAG 1182
Db 518 AACACACAAACAAACACACCCACCCCAACCCCAACACCCCAACACACATATCCAC 459
QY 1183 AGCACCAGCAAAAAGGCCCGCCCAACAGCGGACCAACACACACACAAATGAGATTTC 1242
Db 458 ACCACCATCCCAACACACACACACACCCCAACCAAAACCAACCAACCCCAACCA 399
QY 1243 ACCAGTCTCTCCCCCAACCCCGAGCTCGACTGCACAAAC 1279
Db 398 ACNACCCCAACCAACACACACACACACACACACACACCCACAC 362

RESULT 13
AG131986
LOCUS Pan troglodytes DNA, clone: PTB-144D13.R, genomic survey sequence.
ACCESSION AG131986
VERSION AG131986.1 GI:16661664
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 1626)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .1626
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-144D13.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match 3.9%; Score 71.8; DB 9; Length 1626;
Best Local Similarity 39.7%; Pred. No. 6.3e-09;
Matches 238; Conservative 0; Mismatches 362; Indels 0; Gaps 0;

QY 647 AGAATACAAATTCACAAAGACCAACATGTGCTCGTCCAAATATCTCCACCACTGC 706
Db 156 AAANAAACAAACCAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 215

```

```

QY 707 CCACAGCCCGTCCGGAGATCAAACTCACAAGCACCCCAACTGATGTCACCAAACTCAATA 766
Db 216 CCACCANANAAACNNNAACCAANNAACCCNCCACCCNCCNNAACCAAAACAAAAA 275
QY 767 CCAGGACCCCAAGCAGTGATGATGAGGACCTTCGCAACATCCGGCTCAGGGTCCGGAGAAC 826
Db 276 NANCNCNCCAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 335
QY 827 GAGAACCCCAACACAACTTCTGATCGGTTCACCAAGCAAGGGCTTTTCATCAACAATGCCAC 886
Db 336 NNNNNCCCCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 395
QY 887 CCATCTCCTTACCAACCAACGACGCGCACAGCAAGGAGGAAACACACAAACCAATTCCTCC 946
Db 396 CCCAAACCCCAACACCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 455
QY 947 AAGATGCTGTGACTGAACCTAGACAAACAAATTAACACAACTGCAACACCGTCCATGCCCTCTC 1006
Db 456 CACNACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 515
QY 1007 ATAACACTTACCAATCTTACTAACACACCTCCAAACAACTTCAGCACTCTCTCTG 1066
Db 516 NACACACACCCACNANANAAACCAAAACACAAACCAACAAACAAACAAACAAACAAACAAAC 575
QY 1067 CACATTACAAACACCAACCAATGACACACACAGACACCAATCACTGAAATGAGCAAA 1126
Db 576 CACAACCCCAACACCCCGCCCAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 635
QY 1127 CCAGTGGCCCTCGATACAAACCTGCTCCAAACGGGAAATCCCAACACAGCAAGCA 1186
Db 636 CCCACACCCCGCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCC 695
QY 1187 CCAGCAGCAAAAAGGCGCCGCAACACGCGCAACCAACCAACCAACCAACCAACCAACCAAC 1246
Db 696 ACCNCCNCAACCAACCAACCCCAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 755

RESULT 14
AG382036/c
LOCUS Mus musculus molossinus DNA, clone:MSMg01-192P03.TJ, genomic survey
DEFINITION sequence.
ACCESSION AG382036
VERSION AG382036.1 GI:47993241
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1455)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Teukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Teukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY : pBACe3.6
Vector : pBACe3.6

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:35:29 ; Search time 294.875 Seconds

(without alignments)
10215.808 Million cell updates/sec

Title: US-10-066-506A-3

Perfect score: 1841

Sequence: 1 atgaagaccacatgttctc.....tttgtcttttagttgaattc 1841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTCUS COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1303.4	70.8	2164	3	US-08-760-615-3
2	1303.4	70.8	11460	4	US-09-336-910A-1
3	773.8	42.0	2247	3	US-08-760-615-5
4	526	28.6	2298	4	US-09-650-086A-1
5	525.6	28.5	2172	3	US-08-760-615-1
6	74.6	4.1	152132	4	US-09-949-016-13845
7	74.6	4.1	152145	4	US-09-949-016-12371
8	60.4	3.3	34230	4	US-09-949-016-12052
9	60.4	3.3	128470	4	US-09-949-016-13765
10	57.6	3.1	5511	3	US-08-928-361B-2
11	57.6	3.1	5511	4	US-09-588-995A-2
12	57.6	3.1	7334	3	US-08-928-361B-1
13	57.6	3.1	7334	4	US-09-588-995A-1
14	54.2	2.9	30656	4	US-09-949-016-14613
15	54	2.9	522	4	US-09-949-016-103758
16	53.6	2.9	5163	3	US-08-700-651-1
17	53.6	2.9	5163	3	US-08-928-361B-4
18	53.6	2.9	5163	4	US-09-588-995A-4
19	53.6	2.9	5318	3	US-08-700-651-2
20	53.6	2.9	5318	4	US-08-928-361B-3
21	53.6	2.9	5318	4	US-09-588-995A-3
22	52.6	2.9	271134	4	US-09-949-016-12705
23	52.6	2.9	305491	4	US-09-949-016-17550
24	52.4	2.8	39154	4	US-09-949-016-12384
25	52.4	2.8	39154	4	US-09-949-016-12801
26	52.4	2.8	39443	4	US-09-949-016-14326
27	52.4	2.8	39443	4	US-09-949-016-14327

28	51.6	2.8	114793	4	US-10-148-806-3	Sequence 3, Appli
29	51	2.8	57280	4	US-09-949-016-11796	Sequence 11796, A
30	51	2.8	57280	4	US-09-949-016-12843	Sequence 12843, A
31	51	2.8	57280	4	US-09-949-016-12844	Sequence 12844, A
32	51	2.8	57280	4	US-09-949-016-12846	Sequence 12846, A
33	51	2.8	57280	4	US-09-949-016-13542	Sequence 13542, A
34	51	2.8	57280	4	US-09-949-016-13543	Sequence 13543, A
35	51	2.8	57280	4	US-09-949-016-13544	Sequence 13544, A
36	51	2.8	57280	4	US-09-949-016-13545	Sequence 13545, A
37	51	2.8	57280	4	US-09-949-016-14633	Sequence 14633, A
38	51	2.8	57280	4	US-09-949-016-14634	Sequence 14634, A
39	51	2.8	57280	4	US-09-949-016-14635	Sequence 14635, A
40	51	2.8	57280	4	US-09-949-016-14636	Sequence 14636, A
41	51	2.8	57280	4	US-09-949-016-14637	Sequence 14637, A
42	51	2.8	57280	4	US-09-949-016-14638	Sequence 14638, A
43	51	2.8	57280	4	US-09-949-016-14639	Sequence 14639, A
44	51	2.8	57280	4	US-09-949-016-14640	Sequence 14640, A
45	48.8	2.7	60990	4	US-09-949-016-14080	Sequence 14080, A

ALIGNMENTS

RESULT 1
US-08-760-615-3
; Sequence 3, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2164 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Musoke
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 119..2164
US-08-760-615-3

Query Match		70.8%;	Score 1303.4;	DB 3;	Length 2164;				
Best Local Similarity		99.9%;	Pred. No. 0;						
Matches 1304;		Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	ATGAAGACACATGTTTCTTATCAGTCTTATTAATTCAGGGGCAAAAAATCTCCCG	60						
DB	119	ATGAAGACACATGTTTCTTATCAGTCTTATTAATTCAGGGGCAAAAAATCTCCCG	178						
QY	61	ATTTTAGAGTAGTAGTAATTAATCAACCCCAAAATGCGTATCGTCCGGAACT	120						
DB	179	ATTTTAGAGTAGTAGTAATTAATCAACCCCAAAATGCGTATCGTCCGGAACT	238						
QY	121	CTCCAGAACAGAACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAATGTGCT	180						
DB	239	CTCCAGAACAGAACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAATGTGCT	298						
QY	181	GATTCCTCTTTGGAGGACATCCAGCGATGGGCTTTTCAGGACAGGTGACCTCCCAAGAT	240						
DB	299	GATTCCTCTTTGGAGGACATCCAGCGATGGGCTTTTCAGGACAGGTGACCTCCCAAGAT	358						
QY	241	GTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAGTGTACGGATCCC	300						
DB	359	GTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAGTGTACGGATCCC	418						
QY	301	TCTGGAATACTTCTGTGTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA	360						
DB	419	TCTGGAATACTTCTGTGTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA	478						
QY	361	ACTATCCATCATATTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATGG	420						
DB	479	ACTATCCATCATATTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATGG	538						
QY	421	GGAGCAATTTTCTGTATGATCGCATGCTGCTCCACACATGATGACGAGGCAAGTCTC	480						
DB	539	GGAGCAATTTTCTGTATGATCGCATGCTGCTCCACACATGATGACGAGGCAAGTCTC	598						
QY	481	ACTGAAGGAAACATACAGCTATGATGTCATTAAGACAGTGCACAAAATGATTTCTCG	540						
DB	599	ACTGAAGGAAACATACAGCTATGATGTCATTAAGACAGTGCACAAAATGATTTCTCG	658						
QY	541	CGGCAAGGCAAGGGTACCGTCAATGAATCTGACTTCTACTAATAATATTTGCAAGT	600						
DB	659	CGGCAAGGCAAGGGTACCGTCAATGAATCTGACTTCTACTAATAATATTTGCAAGT	718						
QY	601	AGTAACGGAAACGAAACGATGACATGATGTTTCGGGCTCTTCAAGAAATACAAATCT	660						
DB	719	AGTAACGGAAACGAAACGATGACATGATGTTTCGGGCTCTTCAAGAAATACAAATCT	778						
QY	661	ACAAAGAACCAACATGTGCTCGTCCAAATACCTCCACCACTGCGCCACAGCCGTCGG	720						
DB	779	ACAAAGAACCAACATGTGCTCGTCCAAATACCTCCACCACTGCGCCACAGCCGTCGG	838						
QY	721	GAGATCAAACTCAAGACACCCCACTGATGCCCAAACTCAATACCAACGACCCCAAGC	780						
DB	839	GAGATCAAACTCAAGACACCCCACTGATGCCCAAACTCAATACCAACGACCCCAAGC	898						
QY	781	AGTGATGATGAGGACCTCGCAACATCCGCTCAGGGTCCGGAGAACGAGAACCCACACA	840						
DB	899	AGTGATGATGAGGACCTCGCAACATCCGCTCAGGGTCCGGAGAACGAGAACCCACACA	958						
QY	841	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTATCAACAAATGCGCCACTCCCTCACCA	900						
DB	959	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTATCAACAAATGCGCCACTCCCTCACCA	1018						
QY	901	CAACCAAGCACGCGCACAGAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT	960						
DB	1019	CAACCAAGCACGCGCACAGAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT	1078						
QY	961	GAAGTACACAAAATTAACCAACTGACAAACCGTCCATGCGCCCTCATTAACACTACCACA	1020						
DB	1079	GAAGTACACAAAATTAACCAACTGACAAACCGTCCATGCGCCCTCATTAACACTACCACA	1138						
RESULT 2									
US-09-336-910A-1									
; Sequence 1, Application US/09336910A									
; Patent No. 6517842									
; GENERAL INFORMATION:									
; APPLICANT: United States Army Medical Research Institute of									
; APPLICANT: Infectious Diseases									
; APPLICANT: Hevey, Michael C.									
; APPLICANT: Negley, Diane L.									
; APPLICANT: Pushko, Peter									
; APPLICANT: Smith, Jonathan F.									
; APPLICANT: Schmaljohn, Alan L.									
; TITLE OF INVENTION: Marburg Virus Vaccines									
; FILE REFERENCE: Army 143									
; CURRENT APPLICATION NUMBER: US/09/336,910A									
; PRIOR FILING DATE: 1999-06-21									
; PRIOR APPLICATION NUMBER: US 60/091,403									
; NUMBER OF SEQ ID NOS: 7									
; SOFTWARE: Word, Microsoft Office 97, IBM compatible									
; SEQ ID NO 1									
; LENGTH: 11460									
; TYPE: DNA									
; ORGANISM: Marburg Virus									
US-09-336-910A-1									
Query Match		70.8%;	Score 1303.4;	DB 4;	Length 11460;				
Best Local Similarity		99.9%;	Pred. No. 0;						
Matches 1304;		Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	ATGAAGACACATGTTTCTTATCAGTCTTATTAATTCAGGGGCAAAAAATCTCCCG	60						
DB	5940	ATGAAGACACATGTTTCTTATCAGTCTTATTAATTCAGGGGCAAAAAATCTCCCG	5999						
QY	61	ATTTTAGAGTAGTAGTAATTAATCAACCCCAAAATGCGTATCGTCCGGAACT	120						
DB	6000	ATTTTAGAGTAGTAGTAATTAATCAACCCCAAAATGCGTATCGTCCGGAACT	6059						
QY	121	CTCCAGAACAGAACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAATGTGCT	180						
DB	6060	CTCCAGAACAGAACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAATGTGCT	6119						
QY	181	GATTCCTCTTTGGAGGACATCCAGCGATGGGCTTTTCAGGACAGGTGACCTCCCAAGAT	240						
DB	6120	GATTCCTCTTTGGAGGACATCCAGCGATGGGCTTTTCAGGACAGGTGACCTCCCAAGAT	6179						
QY	241	GTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAGTGTACGGATCCC	300						
DB	6180	GTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAGTGTACGGATCCC	6239						
QY	301	TCTGGAATACTTCTGTGTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA	360						

Db 6240 TCTGGAAATCCTTGCTGTAGATCCTCTACCAACATCCGAGTATCTTAATGCAA 6299
QY 361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTTATGG 420
Db 6300 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTTATGG 6359
QY 421 GGAGCATTTTTCTGTATGATCGCATTCGCTCCCAACAATGTACCGAGGCAATGCTTC 480
Db 6360 GGAGCATTTTTCTGTATGATCGCATTCGCTCCCAACAATGTACCGAGGCAATGCTTC 6419
QY 481 ACTGAAGGAAACATAGCAGCTATGATTTGTAATAGACAGTGCACAAAATGATTTCTTCG 540
Db 6420 ACTGAAGGAAACATAGCAGCTATGATTTGTAATAGACAGTGCACAAAATGATTTCTTCG 6479
QY 541 CGGCAAGGAAACATAGGATCGCTCATATGAAATCTGACTTCTACTAATAATATTTGCAAGT 600
Db 6480 CGGCAAGGAAACATAGGATCGCTCATATGAAATCTGACTTCTACTAATAATATTTGCAAGT 6539
QY 601 AGTAACGGAAACGCAACCAATGACACTGGATGTTTCGGGGCTCTTCAAGAATACAAATCT 660
Db 6540 AGTAACGGAAACGCAACCAATGACACTGGATGTTTCGGGGCTCTTCAAGAATACAAATCT 6599
QY 661 ACAAGAAACCAAAATGCTCGCTCCCAAAATACCTCCACCACTGCCACAGCCCGTCCG 720
Db 6600 ACAAGAAACCAAAATGCTCGCTCCCAAAATACCTCCACCACTGCCACAGCCCGTCCG 6659
QY 721 GAGTCAAACTCAAGCAACCCCAACTGATGCCCAAACTCAATPACCAAGCAACCAAGC 780
Db 6660 GAGTCAAACTCAAGCAACCCCAACTGATGCCCAAACTCAATPACCAAGCAACCAAGC 6719
QY 781 AGTGATGATGAGGACTCGCAACATCGGCTCGAGGTCGGAGACGAGACCCCAACA 840
Db 6720 AGTGATGATGAGGACTCGCAACATCGGCTCGAGGTCGGAGACGAGACCCCAACA 6779
QY 841 ACTTCTGATGCGGTCAACAAGGAGGCTTTTCATCAAAATGCCACCACTCCCTCCACA 900
Db 6780 ACTTCTGATGCGGTCAACAAGGAGGCTTTTCATCAAAATGCCACCACTCCCTCCACA 6839
QY 901 CAACCAAGCAGCCCAAGAGGAGGAAACAACAACCAATGCCACCACTCCCTCCACA 960
Db 6840 CAACCAAGCAGCCCAAGAGGAGGAAACAACAACCAATGCCACCACTCCCTCCACA 6899
QY 961 GNACTAGCAAAATTAACAACACTGCAACCGTTCATCGCCCTCATAACTACCAACA 1020
Db 6900 GNACTAGCAAAATTAACAACACTGCAACCGTTCATCGCCCTCATAACTACCAACA 6959
QY 1021 ATCTCTACTTAAACAACCTCCAAACAACCTTCCAGCACTCTCTGACCACTTACAAAAC 1080
Db 6960 ATCTCTACTTAAACAACCTCCAAACAACCTTCCAGCACTCTCTGACCACTTACAAAAC 7019
QY 1081 ACCACCAATGACACACAGAGCAATCACTGAAAATGAGCAAAACCAAGTCCCTCCG 1140
Db 7020 ACCACCAATGACACACAGAGCAATCACTGAAAATGAGCAAAACCAAGTCCCTCCG 7079
QY 1141 ATAAACAACCTGCTCCCAACGGAAATCCCAACAGCAAGAGACCAAGCAAGCAAAA 1200
Db 7080 ATAAACAACCTGCTCCCAACGGAAATCCCAACAGCAAGAGACCAAGCAAGCAAAA 7139
QY 1201 GSCCCCGCACAAACGGCAACCAACAGCAATGAGCATTTCCAGCACTCTCCCTCCAC 1260
Db 7140 GSCCCCGCACAAACGGCAACCAACAGCAATGAGCATTTCCAGCACTCTCCCTCCAC 7199
QY 1261 CCAGCTCGACTGCAACATCTTGTATATTTTCAAGAAAGCGA 1305
Db 7200 CCAGCTCGACTGCAACATCTTGTATATTTTCAAGAAAGCGA 7244

RESULT 3

US-08-760-615-5
; Sequence 5, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R

; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2247 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Ravn
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..2142
; US-08-760-615-5

Query Match 42.0%; Score 773.8; DB 3; Length 2247;
Best Local Similarity 74.6%; Pred. No. 5e-243;
Matches 973; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
Db 97 ATGAAGACCAATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCC 156
QY 61 ATTTTAGAGTAGCTAGTAAATAATCAACCCCAATGTTGGATTGGTATGCTCCGGAAT 120
Db 157 GTTTTAGAAATTTGTTAGTAACAGCAACCTCAAGATGTAGATTGATGTCTCCGGAAC 216
QY 121 CTCAGAAGACAGAGAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 217 CTCGAAGACAGAGAGTGTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 276
QY 181 GATTCCCTTTGGAGGCAATCAAGCGATCGGCTTTTCAAGACAGGTGTACTCTCCCAAGAT 240
Db 277 GATTCCCTTTGGAAGCATCTAAACGATGGGCTTTTCAAGACAGGTGTACTCTCCCAAGAC 336
QY 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATAAGTAAAGTAAACGATCCC 300
Db 337 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACAATAAGTAAAGTAAACGATCCC 396
QY 301 TCTGGAATAATCTTGTGTTAGATCTCTCTACCAACATCCGCTACTATCCGGAATGCAAA 360
Db 397 TCTGGAATAATCTTGTGTTAGATCTCTCTACCAAGTAAATATCCCGGATTTACCTTAATGTAAA 456

ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53703
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/760,615
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Berson, Bennett J
REGISTRATION NUMBER: 37094
REFERENCE/DOCKET NUMBER: 110229.91241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2172 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Ebola virus
STRAIN: Zaire
FEATURE:
NAME/KEY: CDS
LOCATION: 142..2172
OTHER INFORMATION: /product= "Glycoprotein"
US-08-760-615-1

Query Match 28.5%; Score 525.6; DB 3; Length 2172;
Best Local Similarity 98.3%; Pred. No. 1.9e-161;
Matches 531; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy	1294	AGAAGAAACGATCGGCAATGTGCAATGCTCAACCCAAATGCAACCTTAATTACATTAC	1353
Db	1633	AGAACTCGAAGAGAAGCAATGTGCAATGCTCAACCCAAATGCAACCTTAATTACATTAC	1692
Qy	1354	TGGACTACTCAGGATGAAGGTGCTCAATCGGACTGGCTCGATACCATATTTTCGGGCCA	1413
Db	1693	TGGACTACTCAGGATGAAGGTGCTCAATCGGACTGGCTCGATACCATATTTTCGGGCCA	1752
Qy	1414	GCAGCCGAGGGAATTTACATAGAGGGGCTAATGCACAATCAAGATGGTTTAATCTGTGGG	1473
Db	1753	GCAGCCGAGGGAATTTACATAGAGGGGCTAATGCACAATCAAGATGGTTTAATCTGTGGG	1812
Qy	1474	TTGAGACAGCTGGCCAAAGAGACGCTCAAGCTTTCAACTGTTCTCGAGGCCAACAAT	1533
Db	1813	TTGAGACAGCTGGCCAAAGAGACGCTCAAGCTTTCAACTGTTCTCGAGGCCAACAAT	1872
Qy	1534	GAGCTACCCACCTTTTCAATCTCAACCGTAGGCAATTTGATTTCTGCTCGAGCGATGG	1593
Db	1873	GAGCTACCCACCTTTTCAATCTCAACCGTAGGCAATTTGATTTCTGCTCGAGCGATGG	1932
Qy	1594	GGCGGCACATGCAATTTCTGGGACGGGACTGCTGTATCGAACCAACATGATTTGACCAAG	1653
Db	1933	GGCGGCACATGCAATTTCTGGGACGGGACTGCTGTATCGAACCAACATGATTTGACCAAG	1992
Qy	1654	AACTAAACAGACAAATTTGATCAGATTATCATGATTTTGTGATAAAACCTTCGCGAC	1713
Db	1993	AACTAAACAGACAAATTTGATCAGATTATCATGATTTTGTGATAAAACCTTCGCGAC	2052
Qy	1714	CAGGGGGCAATGCAATTTGGTGGACAGATGGAGCAATGGATACCGCGAGGATTGGA	1773
Db	2053	CAGGGGGCAATGCAATTTGGTGGACAGATGGAGCAATGGATACCGCGAGGATTGGA	2112

Qy	1774	GTTACAGCGGTATATAATTGCAAGTTATCGCTTTATTTCTGTATATGCAAAATTTGCTTTTAG	1833
Db	2113	GTTACAGCGGTATATAATTGCAAGTTATCGCTTTATTTCTGTATATGCAAAATTTGCTTTTAG	2172

RESULT 6
US-09-949-016-13845/c
; Sequence 13845, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13845
; LENGTH: 152132
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13845

Query Match 4.1%; Score 74.6; DB 4; Length 152132;
Best Local Similarity 50.6%; Pred. No. 7.8e-12;
Matches 206; Conservative 0; Mismatches 199; Indels 2; Gaps 1;

Qy	872	CATCAACAATGCCACCCACTCCCTCCACCAACAACCAAGCAGCCACCAAGGAGGAACA	931
Db	136819	CATCACCACCAACCCACCCACCCACCCACCCCTCCACCAACCCACCCACCATCACCA	136760
Qy	932	ACACAAACCATTTCCCAAGATGCTGTGACTGAACCTAGACAAATAACACAACTGCACAAC	991
Db	136759	CACCACCAACCAACCAACCAACCCACCTCCACCAACCAACCAACCTCCACCAACCAAC	136700
Qy	992	CGT--CGATGCCCTCATACCACTACCACAATCTCTACTAACCAACCTCCCAACACAA	1049
Db	136699	CATCACCATCACCAACCCACCCACCCACCCACCCACCATCATCACCACCAACCAAC	136640
Qy	1050	CTTCAGCACTCTCTCTGACGATTAACAAACACCAACCAACCAACCAACCAACCAAC	1109
Db	136639	CACCACCAACCAACCAACCCACCCACCCCTCCACCAACCAACCAACCAACCAACCAAC	136580
Qy	1110	CACGTGAAATGAGCAAAACCAAGTGCCTCGATAAACCAACCTCCCAACGGGAATCC	1169
Db	136579	CACCAACCAACCAACCTTCACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	136520
Qy	1170	CACCACAGCAAGAGCAGCAGCAAGGCGCCGCAACCAACGGGCAACCAACCAACCAAC	1229
Db	136519	CACCACCAACCAACCAACCCACCCACCCACCCACCCACCCACCTACTACCACTACCA	136460
Qy	1230	AAATGAGCATTTTACCAAGTCTCTCCCAACCCCAACCCCAACCCCAACCCCAACCC	1276
Db	136459	CACCATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC	136413

RESULT 7
US-09-949-016-12371/c
; Sequence 12371, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307

```
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12371
; LENGTH: 152145
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12371

Query Match
Best Local Similarity 4.1%; Score 74.6; DB 4; Length 152145;
Matches 206; Conservative 0; Mismatches 199; Indels 2; Gaps 1;

QY 872 CATCAACATGCGCCACCTCCCTCACCACCAACCAAGCAGCGCCACGCAAGGAGGAACA 931
Db 136819 CATCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 136760

QY 932 ACACAAACATCCCAAGATGCTGTGATGAATAGACAAATAACACAACTGCACAC 991
Db 136759 CACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 136700

QY 992 CGT--CCATGCCCTCATACACTACACATCTCTACTACACACCTCCCAACACAA 1049
Db 136699 CATCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 136640

QY 1050 CTTGAGCACTCTCTGCAACCATTAACAAACACCACCACCAATGACACACAGACAA 1109
Db 136639 CACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 136580

QY 1110 CACTGAAATGAGCAACCAAGTCCCTCGATACAAACCTGCTCCCAACGGGAAATCC 1169
Db 136579 CACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 136520

QY 1170 CACCACGAAAGAGCAGCAGCAGCAAAAGGCGCCGACACAGCGGCAACACGAC 1229
Db 136519 CACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 136460

QY 1230 AAATGAGCATTTCCACAGTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1276
Db 136459 CACCATCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 136413

RESULT 8
US-09-949-016-12052
; Sequence 12052, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12052
; LENGTH: 34230
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12052
```

```
Query Match
Best Local Similarity 3.3%; Score 60.4; DB 4; Length 34230;
Matches 136; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 1009 AACACTACCAATCTCTACTAAACAACACCTCCAAACACAACTTCAGCACTCTCTCTGCA 1068
Db 4380 ACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 4439

QY 1069 CCATTACAAAACACACCAATGACACACACAGACGACAAATCACTGAAATGAGCAAAAC 1128
Db 4440 ACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 4499

QY 1129 AGTGCCCTCGATAAACAACCTGCTCCCAAGGGGAAATCCACACAGCAAGAGCACC 1188
Db 4500 ACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 4559

QY 1189 AGCAGCAAAAAGGCGCCGACCAACCGGACCAACAAACAGCAAAATGAGCAATTTACCAAGT 1248
Db 4560 ACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 4619

QY 1249 CCTCCCCCCCCCAGCTCGA 1270
Db 4620 ACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 4641

RESULT 9
US-09-949-016-13765
; Sequence 13765, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13765
; LENGTH: 128470
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13765
```

```
Query Match
Best Local Similarity 3.3%; Score 60.4; DB 4; Length 128470;
Matches 136; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 1009 AACACTACCAATCTCTACTAAACAACACCTCCAAACACAACTTCAGCACTCTCTCTGCA 1068
Db 98325 ACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 98384

QY 1069 CCATTACAAAACACCAATGACACACACAGACGACAAATCACTGAAATGAGCAAAAC 1128
Db 98385 ACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 98444

QY 1129 AGTGCCCTCGATAAACAACCTGCTCCCAAGGGGAAATCCACACAGCAAGAGCACC 1188
Db 98445 ACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 98504

QY 1189 AGCAGCAAAAAGGCGCCGACCAACCGGACCAACAAACAGCAAAATGAGCAATTTACCAAGT 1248
Db 98505 ACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACCACC 98564

QY 1249 CCTCCCCCCCCCAGCTCGA 1270
Db 1249 CCTCCCCCCCCCAGCTCGA 1270
```

Db 98565 ACCACCACCACCACCATCACCA 98586

RESULT 10

US-08-928-361B-2

; Sequence 2, Application US/08928361B

; Patent No. 6071518

; GENERAL INFORMATION:

; APPLICANT: Petersen, Carolyn

; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,

; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS

; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM

; TITLE OF INVENTION: SPECIES INFECTIONS

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PETERS, VERNY, JONES & BIKSA

; STREET: 385 Sherman Avenue, Suite 6

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306-1840

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/928,361B

; FILING DATE: 12-SEP-1997

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/026,062

; FILING DATE: 13-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: VERNY, Hana

; REGISTRATION NUMBER: 30,518

; REFERENCE/DOCKET NUMBER: 480.76-1 (HV)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-324-1677

; TELEFAX: 650-324-1678

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5511 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-928-361B-2

Query Match 3.1%; Score 57.6; DB 3; Length 5511;

Best Local Similarity 45.7%; Pred. No. 2.8e-07;

Matches 201; Conservative 0; Mismatches 239; Indels 0; Gaps 0;

QY 832 CCCACACAACTTCTGATCGGTCCACGAGCGGCTTTTCATCAACAATGCCCCACT 891
Db 895 CCATACACTAAATGTTGGAGTGAACACACAAACAACTACTACTACTACT 954
QY 892 CCCTCACCACAAACGACGCGCACAGCAAGGAGGAAACACAAACCATTTCCCAAGAT 951
Db 955 ACTACTAGCAGACAAACAAACGACGACAAACAACTACTACTACTACTACT 1014
QY 952 GCTGTGACTGAAGTACAAAAAATACAACTGACAAACCGTCCATGCCCCCTCATAC 1011
Db 1015 ACTACTAGCAGACAAACAACTACTACTAGCAGCAACAAACAACTACTACTACT 1074
QY 1012 ACTTACCAATCTTACTTAAGACACCTTCAACACAACTTCCAGCACTCTCTGACCA 1071
Db 1075 ACAACCAACAACTACTAACCCGACGAACTAACCAACCAACCACTACCAGAAACCA 1134
QY 1072 TTACAAAAACACCAATGACAAACACAGAGACCAATCACTGAAATGAGCAACCACT 1131
Db 1135 ACAACCAACAACTACTTAAGACACCTTACTACTACTACTACTACTACTACTACT 1194
QY 1132 GCCCCTCGATAAACACCTTCCAAACGGGAAATCCCAACGAGCAAGGACCAACC 1191
Db 1195 ACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1254

QY 1132 GCCCCTCGATAAACACCTTCCAAACGGGAAATCCCAACGAGCAAGGACCAACC 1191
Db 1195 ACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 1254
QY 1192 AGCAAAAAAGGCCCGGCACAAACGGCACCACCAACGAAATGAGCATTTTCCACGAGTCT 1251
Db 1255 ACAACTACCAAGAAACCAACAAACAAACAAACAAACAACTACTACTACTACTACT 1314
QY 1252 CCCCCCACCCTCCAGCTCGAC 1271
Db 1315 ACGACAACAAACCAACGAC 1334
RESULT 11
US-09-588-995A-2
; Sequence 2, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5
; CURRENT APPLICATION NUMBER: US/09/588,995A
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 08/827,171
; PRIOR FILING DATE: 1997-03-27
; PRIOR APPLICATION NUMBER: 08/928,361
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 08/700,651
; PRIOR FILING DATE: 1996-08-14
; PRIOR APPLICATION NUMBER: 08/415,751
; PRIOR FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2
; LENGTH: 5511
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-09-588-995A-2
Query Match 3.1%; Score 57.6; DB 4; Length 5511;
Best Local Similarity 45.7%; Pred. No. 2.8e-07;
Matches 201; Conservative 0; Mismatches 239; Indels 0; Gaps 0;

Db 1255 ACAACTACCAAGAAACCAACAAACAAACAAACAAACAACTACTACTACAACC 1314
QY 1252 CCCCCCAGGAGTGGAC 1271
Db 1315 ACACACACACACACACGAC 1334

RESULT 12
US-08-928-361B-1
; Sequence 1, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7334 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-928-361B-1

Query Match 3.1%; Score 57.6; DB 3; Length 7334;
Best Local Similarity 45.7%; Pred. No. 3.5e-07;
Matches 201; Conservative 0; Mismatches 239; Indels 0; Gaps 0;

QY 832 CCCCACACACTTCTGATGGGTCCACCAAGCAGGCTTTTCATCAACATGCGCCACT 891
Db 2563 CCATACATAAATGTGTGGAGTGAACACACAAACAACTACTACTACTACT 2622

QY 892 CCTCACCACACCAAGCAGCGCCACGAGGAGGAAACACAAACCATTTCCCAAGAT 951
Db 2623 ACTACTACGACACACACACACACACACACACACACACTACTACTACTACT 2682

QY 952 GTGTGACTGAAGTACGACAAATAACCAACTGCGACACCGTCCATGCCCTCATAC 1011
Db 2683 ACTACTACGACACACACACACACACACACACACACACTACTACTACTACT 2742

QY 1012 ACTACCAACTCTCTACTAACACACCTCCAAACACAACTTCGACACTCTCTGCACCA 1071
Db 2683 ACTACTACGACACACACACTACTACTAGCAACAAACAACTACTACTACTACT 2742

QY 1072 TTACAAAAACACCAACATGACACACACACACACACTCTGAAATGAGCAACACT 1131
Db 2743 ACAACCAACAACTACTACCAACCAACCAACCAACCAACCAACCAACCAACCA 2802

QY 1072 TTACAAAAACACCAACATGACACACACACACACACTGAAAAATGAGCAACACT 1131
Db 2803 ACAACCAACAAACAAACAACTACTACTACTACTACTACTACTACTACTACTACT 2862

QY 1132 GCGCCCTGATAACAAACCTCCCAAGGGAATCCACACACAGCAGCAACGAGCAGC 1191
Db 2863 ACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 2922

QY 1192 AGCAAAAAAGGCGCCGACACAAACGACCAACCAACCAACCAACCAACCAACCA 1251
Db 2923 ACACTACCAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2982

QY 1252 CCCCCCAGGAGTGGAC 1271
Db 2983 ACACACACACACACACGAC 3002

RESULT 13
US-09-588-995A-1
; Sequence 1, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5
; CURRENT APPLICATION NUMBER: US/09/588,995A
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 08/827,171
; PRIOR FILING DATE: 1997-03-27
; PRIOR APPLICATION NUMBER: 08/928,361
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 08/700,651
; PRIOR FILING DATE: 1996-08-14
; PRIOR APPLICATION NUMBER: 08/415,751
; PRIOR FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7334
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-09-588-995A-1

Query Match 3.1%; Score 57.6; DB 4; Length 7334;
Best Local Similarity 45.7%; Pred. No. 3.5e-07;
Matches 201; Conservative 0; Mismatches 239; Indels 0; Gaps 0;

QY 832 CCCCACACACTTCTGATGGGTCCACCAAGCAGGCTTTTCATCAACATGCGCCACT 891
Db 2563 CCATACATAAATGTGTGGAGTGAACACACAAACAACTACTACTACTACTACT 2622

QY 892 CCTCACCACACCAAGCAGCGCCACGAGGAGGAAACACAAACCATTTCCCAAGAT 951
Db 2623 ACTACTACGACACACACACACACACACACACACACACTACTACTACTACTACT 2682

QY 952 GTGTGACTGAAGTACGACAAATAACCAACTGCGACACCGTCCATGCCCTCATAC 1011
Db 2683 ACTACTACGACACACACACACACACACACACACACACTACTACTACTACTACT 2742

QY 1012 ACTACCAACTCTCTACTAACACACCTCCAAACACAACTTCGACACTCTCTGCACCA 1071
Db 2743 ACAACCAACAACTACTACCAACCAACCAACCAACCAACCAACCAACCAACCA 2802

QY 1072 TTACAAAAACACCAACATGACACACACACACACTCTGAAATGAGCAACACT 1131
Db 2803 ACAACCAACAACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 2862

```
QY 1132 GCCCCTCGATAACACCTCGCTCCAAACGGGAATCCACACAGCAAGAGCAGCAGC 1191
Db 2863 ACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 2922
QY 1192 AGCAAAAAGGCCCGCCACACAGCGCACAAACAGCAAAATGACATTTTCAACCAAGTCT 1251
Db 2923 ACAACTACCAAGAAACCAACACACAAACAAACAAACAAACAACTACTACTACTAC 2982
QY 1252 CCCCCCAGCCCGAGTCGAC 1271
Db 2983 ACGACAACAACAACACGAC 3002

RESULT 14
US-09-949-016-14613/c
; Sequence 14613, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14613
; LENGTH: 30656
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(30656)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14613

Query Match 2.9%; Score 54.2; DB 4; Length 30656;
Best Local Similarity 14.5%; Pred. No. 1.3e-05;
Matches 164; Conservative 0; Mismatches 966; Indels 0; Gaps 0;

QY 142 CATCTGATGGGATTACACTGAGTGGCAAAAGTTGCTGATTCCTTTTGAGGACATCC 201
Db 4633 CAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4574
QY 202 AAGCGATGGCTTTTCAGGACAGGTGTACTCTCCAAAGAAATGTTGAGTACACAGAGGGGAG 261
Db 4573 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4514
QY 262 GAAGCAAAACATGTACATATAGTGAACGGATCCCTCTGGAAAAATCCTGCTGTGTA 321
Db 4513 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4454
QY 322 GATCTCTTACCAACATCGTCTACTATCGAAATGCAAAACTATCCATCATATTCAAGT 381
Db 4453 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4394
QY 382 CAAAAACCTTCATGACAGGGGATCGCCCTTCATTATGGGGAGCATTTTTCGTGTATGAT 441
Db 4393 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4334
QY 442 CGATTGCTCCACAACAATGTACGAGGCAAGTCTTCACTGAAGGGAAACATAGCAGCT 501
Db 4333 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4274
QY 502 ATGATTGTCAATAGACAGTGCACAAAATGATTTCTCGCGCAAGGACAGGATACCGT 561
Db 4273 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4214
```

```
QY 562 CATATGATCTGACTTCTACTATAATAATTGGACAAGTAGTAACGGAACGCAACGAT 621
Db 4213 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4154
QY 622 GACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAAATCTACAAAGAAACAAATGTGCT 681
Db 4153 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4094
QY 682 CCGTCCAAAAATACCTCCACCACTGCCCAAGCCGTCGGGAGATCAAACTCAACAGCACC 741
Db 4093 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4034
QY 742 CCAACTGATGCCACCAAACTCAATACCAAGCAACCAAGCAGTGTGATGAGGACTCGCA 801
Db 4033 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3974
QY 802 ACATCCGCTCAGGGTCCGGAGAACGAGAACCCCAACCACTTCTGTATGCGGTACCAAG 861
Db 3973 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3914
QY 862 CAAGGGCTTTTATCAACAATGCCACCCCTCTCACCACAACCAAGCAGCAGCAAG 921
Db 3913 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3854
QY 922 GGAGGAAACAAACAAACCAATCCCAAGATGCTGTGACTGAACTAGACAAAAATACACA 981
Db 3853 NNNNNNAACCACTACCAACCAACCAACCAATCCCAAGTACCACTACCAACCAACCA 3794
QY 982 ACTGCACAACCGTCCATGCGCCCTCATACCACTACCACTACCACTACCACTACCACT 1041
Db 3793 ATCACCAGCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCACTACCA 3734
QY 1042 AAACACAACCTTCAGCACTCTCTGTGACCACTTACAAAACCAACCAACCAACCAAC 1101
Db 3733 AGGACCATTACCATCATCTACCACTACCACTACCACTACCACTACCACTACCACTACA 3674
QY 1102 AGCAATCACTGAAATGAGCAAAACGAGTCCCTCGTGTGATTAACAACTCTGCTTCAAG 1161
Db 3673 ACCACCACTACCACTACCACTACCACTACCACTACCACTACCACTACCACTACCACT 3614
QY 1162 GGAATCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1221
Db 3613 ACCATCCACAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCACTACCA 3554
QY 1222 AACACGCAAAATGAGCATTTTACCACTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCC 1271
Db 3553 ATCACCACCACTACCACTACCACTACCACTACCACTACCACTACCACTACCACTACCA 3504
```

RESULT 15

```
US-09-949-016-103758/c
; Sequence 103758, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103758
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Human
```

US-09-949-016-103758

```
Query Match      2.9%; Score 54; DB 4; Length 522;
Best Local Similarity 48.7%; Pred. No. 8e-07;
Matches 147; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 969 CAAAATAACACAACCTGGCACAACCGTCCATGCCCCCTCATTAACACTTACCAATCTCTAC 1028
Db 487 CATCACCACCACCAATCACCAGCACCACCAACCAACCAACCAACCAACCAAT 428

QY 1029 TAACAACACCTCCAACACAACCTTCAGCACTCTCTGCAACCAATTACAAAACCAACCA 1088
Db 427 CACCAGCACCACCATCACCACCAACCACTTACCATTACCAACCAATCACAATCACCAGCAC 368

QY 1089 TGACAACACACAGAGCACAATCACTGMAAATGAGCAAAACCACTGCCCCCTCGATAACAAC 1148
Db 367 CAGACCAACACCAATCAGAACCACTACCAACCACTACCACTACCACTACCACTACCACTAC 308

QY 1149 CCTGCTTCCAACGGGAAATCCCAACACAGCAAGAGCAGCAGCAAAAGGCCCCGC 1208
Db 307 CATTACSAACCATCACTTACCACCATCACCACAATCACCAGCAGCAGCAGCAGCAGCAGCAG 248

QY 1209 CACAAGGGCACCACAACAGCAGCAAAATGAGCATTTTCAACAGTCTCTCCCCCAGCTC 1268
Db 247 CACAACCACTTACCACCATCACCACCACTTACCATCATCATCATCATCATCATCATCATCAT 188

QY 1269 GA 1270
Db 187 CA 186
```

Search completed: August 1, 2005, 10:31:08
Job time : 298.875 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 16:57:35 ; Search time 1157.01 Seconds

(without alignments)
10298.647 Million cell updates/sec

Title: US-10-066-506A-3

Perfect score: 1841

Sequence: 1 atgaagaccacatgttctct.....ttgtcttttagtgaattc 1841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1841	100.0	1841	15	US-10-066-506A-3
2	1308	71.0	2046	15	US-10-066-506A-5
3	1303.4	70.8	7005	20	US-10-491-121-31
4	1303.4	70.8	7778	20	US-10-491-121-30
5	1303.4	70.8	8256	20	US-10-491-121-32
6	1303.4	70.8	11460	16	US-10-267-322-1
7	1274	69.2	2051	15	US-10-066-506A-9

8	1148.2	62.4	2046	21	US-10-811-353-2	Sequence 2, Appli
9	1148.2	62.4	19112	17	US-10-353-856-27	Sequence 27, Appl
10	780	42.4	2046	15	US-10-066-506A-7	Sequence 7, Appl
11	780	42.4	2046	15	US-10-066-506A-13	Sequence 13, Appl
12	688	37.4	6902	20	US-10-491-121-42	Sequence 42, Appl
13	540	29.3	2039	15	US-10-066-506A-11	Sequence 11, Appl
14	527.6	28.7	7272	22	US-10-860-878-4	Sequence 4, Appl
15	527.6	28.7	7285	22	US-10-860-878-3	Sequence 3, Appl
16	527.6	28.7	18959	17	US-10-353-856-19	Sequence 19, Appl
17	527.6	28.7	18959	17	US-10-353-856-37	Sequence 37, Appl
18	527.6	28.7	18959	17	US-10-353-856-46	Sequence 46, Appl
19	526	28.6	2298	9	US-09-337-946A-1	Sequence 1, Appl
20	526	28.6	2298	17	US-10-384-976-1	Sequence 1, Appl
21	526	28.6	2298	18	US-10-226-795-1	Sequence 1, Appl
22	526	28.6	2298	22	US-10-696-633-1	Sequence 10, Appl
23	526	28.6	6467	20	US-10-491-121-10	Sequence 3, Appl
24	526	28.6	6624	20	US-10-491-121-3	Sequence 2, Appl
25	526	28.6	7154	20	US-10-491-121-1	Sequence 2, Appl
26	526	28.6	7188	20	US-10-491-121-2	Sequence 28, Appl
27	526	28.6	8439	20	US-10-491-121-28	Sequence 4, Appl
28	525.6	28.5	6561	20	US-10-491-121-4	Sequence 4, Appl
29	514.4	27.9	2030	21	US-10-811-353-4	Sequence 9, Appl
30	451.8	24.5	6914	20	US-10-491-121-9	Sequence 29, Appl
31	451.8	24.5	8199	20	US-10-491-121-29	Sequence 25, Appl
32	451.8	24.5	10783	20	US-10-491-121-25	Sequence 8, Appl
33	420	22.8	7106	20	US-10-491-121-7	Sequence 7, Appl
34	336.2	18.3	7044	20	US-10-491-121-16	Sequence 16, Appl
35	324.6	17.6	7002	20	US-10-491-121-21	Sequence 21, Appl
36	324.6	17.6	7023	20	US-10-491-121-17	Sequence 17, Appl
37	324.6	17.6	7036	20	US-10-491-121-1	Sequence 1, Appl
38	324.2	17.6	7003	22	US-10-860-878-1	Sequence 37, Appl
39	290.4	15.8	6324	20	US-10-491-121-37	Sequence 12, Appl
40	286.6	15.6	2360	17	US-10-397-635-12	Sequence 19, Appl
41	285	15.5	6889	20	US-10-491-121-19	Sequence 20, Appl
42	285	15.5	8146	20	US-10-491-121-20	Sequence 9, Appl
43	283.4	15.4	18890	17	US-10-353-856-9	Sequence 2, Appl
44	271.6	14.8	7073	22	US-10-860-878-2	Sequence 13, Appl
45	271.6	14.8	7082	20	US-10-491-121-13	

ALIGNMENTS

RESULT 1

US-10-066-506A-3
; Sequence 3, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 3
; LENGTH: 1841
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein 2
; OTHER INFORMATION: Ebola virus Zaire Mayinga strain Glycoprotein 2
US-10-066-506A-3

Query Match 100.0%; Score 1841; DB 15; Length 1841;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAAGACCACATGTTTCCTTATCATCTTATCTATCAAGGACAAAATCTCCC 60

```
Db 1 ATGAAGACACATGTTTCCTTATCAGTCTTAAATCAAGGGACAAAAATCTCCCC 60
QY 61 ATTTTAGAGTAGCTAGTAATCAACCCCAAAATGTGGATTCGTTATGCTCCGGAACT 120
Db 61 ATTTTAGAGTAGCTAGTAATCAACCCCAAAATGTGGATTCGTTATGCTCCGGAACT 120
QY 121 CTCGAGAGCAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCGAGAGCAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCTCTTTGGAGGATCCGAGGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAAAT 240
Db 181 GATTCCTCTTTGGAGGATCCGAGGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAAAT 240
QY 241 GTTGTAGTACACAGAGGGGAGAGCCCAAAACATGCTACAATATAAGTGTACGGATCCC 300
Db 241 GTTGTAGTACACAGAGGGGAGAGCCCAAAACATGCTACAATATAAGTGTACGGATCCC 300
QY 301 TCTGGAATAATCCTTGTCTGTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA 360
Db 301 TCTGGAATAATCCTTGTCTGTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA 360
QY 361 ACTATCCATCATATTCGAGGTCAAAACCTCATGACAGGGGATCCGCTTCATTTATGG 420
Db 361 ACTATCCATCATATTCGAGGTCAAAACCTCATGACAGGGGATCCGCTTCATTTATGG 420
QY 421 GGAGCATTTTTTCTGTATGATCGCATGCTCCACACAACTGTACCGAGGCAAGCTCTTC 480
Db 421 GGAGCATTTTTTCTGTATGATCGCATGCTCCACAACTGTACCGAGGCAAGCTCTTC 480
QY 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTCTCG 540
Db 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTCTCG 540
QY 541 CGGCAAGGACAAAGGTPACCGTCAATGAAATCTGACTTCTACTAATAAATATTTGGCAAGT 600
Db 541 CGGCAAGGACAAAGGTPACCGTCAATGAAATCTGACTTCTACTAATAAATATTTGGCAAGT 600
QY 601 AGTACGGAACGCAACGATGACATGCTGCTCCACAACTGTACCGGCTCTTCAGAAATACAAATCT 660
Db 601 AGTACGGAACGCAACGATGACATGCTGCTCCACAACTGTACCGGCTCTTCAGAAATACAAATCT 660
QY 661 ACAAGAACCAAAACATGCTCGTCCAAATACCTCCACCACTGCCACACGCCGCTCG 720
Db 661 ACAAGAACCAAAACATGCTCGTCCAAATACCTCCACCACTGCCACACGCCGCTCG 720
QY 721 GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCAATACCAAGGACCCCAAGC 780
Db 721 GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCAATACCAAGGACCCCAAGC 780
QY 781 AGTGTATGAGGACCTCGCAACNTCCGCTCAGGGTCCGGAGNAGAGAACCCACACA 840
Db 781 AGTGTATGAGGACCTCGCAACNTCCGCTCAGGGTCCGGAGNAGAGAACCCACACA 840
QY 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTATCAACAAATGCCACCACTCCCTCACCA 900
Db 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTATCAACAAATGCCACCACTCCCTCACCA 900
QY 901 CAACCAAGCAGCAGCAGAGGAGGAAACACACAAACCAATTCCTCAAGATGCTGTGACT 960
Db 901 CAACCAAGCAGCAGCAGAGGAGGAAACACACAAACCAATTCCTCAAGATGCTGTGACT 960
QY 961 GAACTAGACAAATTAACCACTGACACACCGTCCATGCTCCCTCATTAACACTACCA 1020
Db 961 GAACTAGACAAATTAACCACTGACACACCGTCCATGCTCCCTCATTAACACTACCA 1020
QY 1021 ATCTCTACTAAACACCTCCAAACCAACTTCAGACTCTCTCTGCACCAATTAACAAAC 1080
Db 1021 ATCTCTACTAAACACCTCCAAACCAACTTCAGACTCTCTCTGCACCAATTAACAAAC 1080
QY 1081 ACCACCAATGACAAACACAGAGCACAATCACTGAAAAATGAGCAAAACCAAGTCCCTCG 1140
```

```
Db 1081 ACCACCAATGACAAACACAGAGCACAATCACTGAAAAATGAGAAAAACAGTCCCTCG 1140
QY 1141 ATAAACAACCTGCTCCCAACGCGAAATCCCAACAGCAAAAGAGACACAGAGCAAAAAA 1200
Db 1141 ATAAACAACCTGCTCCCAACGCGAAATCCCAACAGCAAAAGAGACACAGAGCAAAAAA 1200
QY 1201 GGCCCCGCCACAAGGGCACCAACACGACAAATGAGCAATTTCAACAGTCTCTCCCTCC 1260
Db 1201 GGCCCCGCCACAAGGGCACCAACACGACAAATGAGCAATTTCAACAGTCTCTCCCTCC 1260
QY 1261 CCCAGCTCGACTGCACACATCTTGTATATTTTTCAGAGAAAGCGATCGGCAATTTGCAAT 1320
Db 1261 CCCAGCTCGACTGCACACATCTTGTATATTTTTCAGAGAAAGCGATCGGCAATTTGCAAT 1320
QY 1321 GCTCAACCAATGCAACCTTAATTTTACATTTACTGGAATCTACTCAGGATGAAGTGTGCA 1380
Db 1321 GCTCAACCAATGCAACCTTAATTTTACTTACTGGAATCTACTCAGGATGAAGTGTGCA 1380
QY 1381 ATCGGACTGGCTGGATACCATATTTTCGGGCGAGCAGCGAGGGAATTTACATAGAGGG 1440
Db 1381 ATCGGACTGGCTGGATACCATATTTTCGGGCGAGCAGCGAGGGAATTTACATAGAGGG 1440
QY 1441 CTAATGCAATCAAGATGGTTTAACTCTGTGGTTTGACACAGCTGCCCAACGAGAGCT 1500
Db 1441 CTAATGCAATCAAGATGGTTTAACTCTGTGGTTTGACACAGCTGCCCAACGAGAGCT 1500
QY 1501 CAAGCTCTTCAACTGCTTCGAGAGCCAACTGAGCTACGACCTTTTCAATCTCTCAAC 1560
Db 1501 CAAGCTCTTCAACTGCTTCGAGAGCCAACTGAGCTACGACCTTTTCAATCTCTCAAC 1560
QY 1561 CGTAAGGCAATGATTTCTTGTGCGAGATGGGCGGCGACATGCCAATTTCTGGGACCG 1620
Db 1561 CGTAAGGCAATGATTTCTTGTGCGAGATGGGCGGCGACATGCCAATTTCTGGGACCG 1620
QY 1621 GACTGCTGTATCGAACCATGATTTGGACCAAGAACATTAACAGAAATTTGATCAGATT 1680
Db 1621 GACTGCTGTATCGAACCATGATTTGGACCAAGAACATTAACAGAAATTTGATCAGATT 1680
QY 1681 ATTCATGATTTTGTGATAAAACCTTCGGACACGAGGGGACAAATGACAAATTTGGTGGACA 1740
Db 1681 ATTCATGATTTTGTGATAAAACCTTCGGACACGAGGGGACAAATGACAAATTTGGTGGACA 1740
QY 1741 GGATGAGACAAATGATACCGGCGAGTATTTGAGTTTACAGGCGTTTAAATTCAGATTATC 1800
Db 1741 GGATGAGACAAATGATACCGGCGAGTATTTGAGTTTACAGGCGTTTAAATTCAGATTATC 1800
QY 1801 GCTTTATTTCTGATATGCAAAATTTGCTCTTTTGTGAAATTC 1841
Db 1801 GCTTTATTTCTGATATGCAAAATTTGCTCTTTTGTGAAATTC 1841
```

RESULT 2

US-10-066-506A-5

; Sequence 5, Application US/10066506A

; Publication No. US20030108560A1

; GENERAL INFORMATION:

; APPLICANT: Grogan, Case C.

; APPLICANT: Hevey, Michael C.

; APPLICANT: Schmaljohn, Alan, L.

; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein

; FILE REFERENCE: 003/243/SAP

; CURRENT APPLICATION NUMBER: US/10/066,506A

; CURRENT FILING DATE: 2002-01-31

; PRIOR APPLICATION NUMBER: 60/267,522

; PRIOR FILING DATE: 2001-01-31

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Apple Macintosh Microsoft Word 6.0

; SEQ ID NO 5

; LENGTH: 2046

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein

OTHER INFORMATION: Marburg virus strain Raven Glycoprotein 2
US-10-066-506A-5

Query Match 71.0%; Score 1308; DB 15; Length 2046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAAGACCATGTTTCTTATGCTTCTTATTAATTAAGGACCAAAATCTCC	60
DB	1	ATGAAGACCATGTTTCTTATGCTTCTTATTAATTAAGGACCAAAATCTCC	60
QY	61	ATTTTAGAGATAGCTAGTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGA	120
DB	61	ATTTTAGAGATAGCTAGTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGA	120
QY	121	CTCCAGAGACAGAAAGAGCTCATCTGATGGGATTCACATGAGTGGGCAAAAATGTTG	180
DB	121	CTCCAGAGACAGAAAGAGCTCATCTGATGGGATTCACATGAGTGGGCAAAAATGTTG	180
QY	181	GATTCCTCTTGGAGGATCCAGCGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT	240
DB	181	GATTCCTCTTGGAGGATCCAGCGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT	240
QY	241	GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGAACGATCCC	300
DB	241	GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGAACGATCCC	300
QY	301	TCTGAAAATCTTCTGTTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA	360
DB	301	TCTGAAAATCTTCTGTTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA	360
QY	361	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTTATGG	420
DB	361	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTTATGG	420
QY	421	GGAGCATTTTCTGTATGATCGATTCCTCCACAAATGTACCGAGGCAAAATCTTC	480
DB	421	GGAGCATTTTCTGTATGATCGATTCCTCCACAAATGTACCGAGGCAAAATCTTC	480
QY	481	ACTGAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCG	540
DB	481	ACTGAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCG	540
QY	541	CGGCAAGACAGAGGTACGTCATATGATCTGACTTCTACTAATAAATATTTGCAAGT	600
DB	541	CGGCAAGACAGAGGTACGTCATATGATCTGACTTCTACTAATAAATATTTGCAAGT	600
QY	601	AGTAAACGGAACCAACGAATGACACTGATGTTTTCGGGCTCTTCAAGAAATACAAATCT	660
DB	601	AGTAAACGGAACCAACGAATGACACTGATGTTTTCGGGCTCTTCAAGAAATACAAATCT	660
QY	661	ACAAAGAACCAACATGTCCTCGTCCAAATATACCTCCACCACTGCCACAGCCCGTCCG	720
DB	661	ACAAAGAACCAACATGTCCTCGTCCAAATATACCTCCACCACTGCCACAGCCCGTCCG	720
QY	721	GAGATCAAACTCAAGACACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC	780
DB	721	GAGATCAAACTCAAGACACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC	780
QY	781	AGTGATGATGAGGACTCGACATCTCCGCTCAGGTCGGGAGACGAGAACCCACACA	840
DB	781	AGTGATGATGAGGACTCGACATCTCCGCTCAGGTCGGGAGACGAGAACCCACACA	840
QY	841	ACTTCTGTGCGGTCAACAGCAAGGGCTTTTCATCAAAATGACCACTCCCTCCACCA	900
DB	841	ACTTCTGTGCGGTCAACAGCAAGGGCTTTTCATCAAAATGACCACTCCCTCCACCA	900
QY	901	CAACCAAGCAACCCACAGCAAGGAGAAACAAACAAACCAATTCCTCCAGATCTGTGAT	960
DB	901	CAACCAAGCAACCCACAGCAAGGAGAAACAAACAAACCAATTCCTCCAGATCTGTGAT	960
QY	961	GNACTAGACAAAATAACACATCTGCACACCGTTCATGCCCCCTCATAACTACCA	1020

DB	961	GAACTAGACAAAATAACACAACTGCACACCGTCCATGCCCCCTCATAACTACCACA	1020
QY	1021	ATCTCTACTAACAACACCTCCAAACACACACTTTCAGCAGCTCTCTCTGCAACATTCACAAAC	1080
DB	1021	ATCTCTACTAACAACACCTCCAAACACACACTTTCAGCAGCTCTCTCTGCAACATTCACAAAC	1080
QY	1081	ACCACCAATGACACACAGAGCAGCAATCACTGAAAATGAGCAAAACCCAGTGCCCTCG	1140
DB	1081	ACCACCAATGACACACAGAGCAGCAATCACTGAAAATGAGCAAAACCCAGTGCCCTCG	1140
QY	1141	ATAACAAACCTCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAGCAGCAAAAAA	1200
DB	1141	ATAACAAACCTCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAGCAGCAAAAAA	1200
QY	1201	GGCCCCGCCACAAACGGGACCAACACAGCAAAATGAGCAATTCACAGTCTCTCCCCCACC	1260
DB	1201	GGCCCCGCCACAAACGGGACCAACACAGCAAAATGAGCAATTCACAGTCTCTCCCCCACC	1260
QY	1261	CCCAGCTCGACTGCACACATCTTTGTATATTTTCAGAGAAAGCGATCG	1308
DB	1261	CCCAGCTCGACTGCACACATCTTTGTATATTTTCAGAGAAAGCGATCG	1308

RESULT 3
US-10-491-121-31
; Sequence 31, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; FILE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 7005
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012x/s Marburg GP(dTM)
US-10-491-121-31

Query Match 70.8%; Score 1303.4; DB 20; Length 7005;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	ATGAAGACCAATGTTTCTTATGCTTATCAGTCTTATCTTAATTAAGGACCAAAAAATCTCCC	60
DB	2034	ATGAAGACCAATGTTTCTTATGCTTATCAGTCTTATCTTAATTAAGGACCAAAAAATCTCCC	2093
QY	61	ATTTTAGAGATAGCTAGTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGA	120
DB	2094	ATTTTAGAGATAGCTAGTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGA	2153
QY	121	CTCCAGAGACAGAAAGAGCTCATCTGATGGGATTCACACTGAGTGGGCAAAAATGTTG	180
DB	2154	CTCCAGAGACAGAAAGAGCTCATCTGATGGGATTCACACTGAGTGGGCAAAAATGTTG	2213
QY	181	GATTCCTCTTGGAGGATCCAGCGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT	240
DB	2214	GATTCCTCTTGGAGGATCCAGCGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT	2273
QY	241	GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGAACGATCCC	300


```
Db 2274 GTTGAGTACACAGAGGGGAGGAAACCAAAACATGCTACAATATAGTGTAACGATCCC 2333
QY 301 TCTGGAATACTTGTGTTAGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA 360
Db 2334 TCTGGNAATCTTGTGTTAGATCCTCTACCAACATCCGTGACTATCCTAAATGCAAA 2393
QY 361 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
Db 2394 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 2453
QY 421 GGAGCATTTTTCTGTATGATCGCATGCTCCACCAACATGACGGGCAAAATCTTC 480
Db 2454 GGAGCATTTTTCTGTATGATCGCATGCTCCACCAACATGACGGGCAAAATCTTC 2513
QY 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTTCG 540
Db 2514 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTTCG 2573
QY 541 CGGCAAGGAAACATAGGCTCATATGAATCTGACTTTCTACTAATAATATTTGGCAAGT 600
Db 2574 CGGCAAGGAAACATAGGCTCATATGAATCTGACTTTCTACTAATAATATTTGGCAAGT 2633
QY 601 AGTAACGGAACCGAAACGATGACACTGATGTTTGGCGCTCTTCAAGAAATACAAATCT 660
Db 2634 AGTAACGGAACCGAAACGATGATGTTTGGCGCTCTTCAAGAAATACAAATCT 2693
QY 661 ACAAGAACCAACATGCTCGTCCAAATACCTCCACCACTGCCACAGCCCGTCG 720
Db 2694 ACAAGAACCAACATGCTCGTCCAAATACCTCCACCACTGCCACAGCCCGTCG 2753
QY 721 GAGATCAAACTCACAAGCACCCCACTGATGCCACCAAACTCAATACCGAGCCCAAGC 780
Db 2754 GAGATCAAACTCACAAGCACCCCACTGATGCCACCAAACTCAATACCGAGCCCAAGC 2813
QY 781 AGTGATGATGAGACCTCCGCAACATCCGCTCAGGCTCGGAGAGAGGAAACCCACACA 840
Db 2814 AGTGATGATGAGAGCTCCGCAACATCCGCTCAGGCTCGGAGAGAGGAAACCCACACA 2873
QY 841 ACTTCTGATGCGGTACCAAGCAAGGCTTTTCAATCAAACTGCCACCACTCCCTCACA 900
Db 2874 ACTTCTGATGCGGTACCAAGCAAGGCTTTTCAATCAAACTGCCACCACTCCCTCACA 2933
QY 901 CAAACCAAGCACGCGCACAGAGGAGGAAACAAACAAACCATTTCCCAAGATCTGTGACT 960
Db 2934 CAAACCAAGCACGCGCACAGAGGAGGAAACAAACAAACCATTTCCCAAGATCTGTGACT 2993
QY 961 GAACTAGACAAAATTAACACAACTGACAAACCGTCCATGCCCTCATTAACACTACCACA 1020
Db 2994 GAACTAGACAAAATTAACACAACTGACAAACCGTCCATGCCCTCATTAACACTACCACA 3053
QY 1021 ATCTCTACTAAACACACCTCCAAACAACTTTCAGCACTCTCTGSCACCATTAACAAAC 1080
Db 3054 ATCTCTACTAAACACACCTCCAAACAACTTTCAGCACTCTCTGSCACCATTAACAAAC 3113
QY 1081 ACCACCAATGACAAACACAGAGGACCAATCACTGAAAATGAGCAAAACCAAGTCCCCCTCG 1140
Db 3114 ACCACCAATGACAAACACAGAGGACCAATCACTGAAAATGAGCAAAACCAAGTCCCCCTCG 3173
QY 1141 ATAAACAACTGCTCCAAACGGGAAATCCCAACACAGGAAAGAGCAAGGAGGAGGAA 1200
Db 3174 ATAAACAACTGCTCCAAACGGGAAATCCCAACACAGGAAAGAGCAAGGAGGAGGAA 3233
QY 1201 GGCCCGCCACAAACGAGCAACAAACAGCAAAATGAGCAATTTCCACAGTCTCCCCCAAC 1260
Db 3234 GGCCCGCCACAAACGAGCAACAAACAGCAAAATGAGCAATTTCCACAGTCTCCCCCAAC 3293
QY 1261 CCAGCTCGAGTGCACAAACATCTTGATATTTTTCAGAGGAAAGCGA 1305
Db 3294 CCAGCTCGAGTGCACAAACATCTTGATATTTTTCAGAGGAAAGCGA 3338
```

RESULT 4

US-10-491-121-30

```
; Sequence 30, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 7778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012 Marburg
US-10-491-121-30
```

```
Query Match 70.8%; Score 1303.4; DB 20; Length 7778;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTCTTATCAGTCTTATCTTTAATTCAGGGACAAAAATCTCCCC 60
Db 2034 ATGAAGACCAATGTTCTTATCAGTCTTATCTTTAATTCAGGGACAAAAATCTCCCC 2093
QY 61 ATTTAGAGTAGCTAGTAATAATCAACCCCAAAATGGATTCCGTATGTCGGAACT 120
Db 2094 ATTTAGAGTAGCTAGTAATAATCAACCCCAAAATGGATTCCGTATGTCGGAACT 2153
QY 121 CTCAGAGACAGAGAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT 180
Db 2154 CTCAGAGACAGAGAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT 2213
QY 181 GATTCCCTTTGGAGGCAATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAA 240
Db 2214 GATTCCCTTTGGAGGCAATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAA 2273
QY 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACNATATAAGTGAACGGATCCC 300
Db 2274 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACNATATAAGTGAACGGATCCC 2333
QY 301 TCTGAAAATCCTTCTGTTAGATCCTCTACCAACATCCGTCGACTATCCGAAATGCAAA 360
Db 2334 TCTGAAAATCCTTCTGTTAGATCCTCTACCAACATCCGTCGACTATCTTAATGCAAA 2393
QY 361 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
Db 2394 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 2453
QY 421 GGAGCATTTTTCTGTATGATCGCATGCTCCACCAACATGACGGGCAAAATCTTC 480
Db 2454 GGAGCATTTTTCTGTATGATCGCATGCTCCACCAACATGACGGGCAAAATCTTC 2513
QY 481 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTTCG 540
Db 2514 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAGACAGTGCACAAAATGATTTTCG 2573
QY 541 CGGCAAGGAAACATAGGCTCATATGAATCTGACTTTCTACTAATAATATTTGGCAAGT 600
Db 2574 CGGCAAGGAAACATAGGCTCATATGAATCTGACTTTCTACTAATAATATTTGGCAAGT 2633
QY 601 AGTAACGGAACCGAAACGATGACACTGATGTTTGGCGCTCTTCAAGAAATACAAATCT 660
Db 2634 AGTAACGGAACCGAAACGATGACACTGATGTTTGGCGCTCTTCAAGAAATACAAATCT 2693
```

```
QY 661 ACAAGAACCAAAACATGTCTCGTCCAAATAATACCTCCACCACTGCGCCACAGCCGCTCG 720
Db |||||
2694 ACAAGAACCAAAACATGTCTCGTCCAAATAATACCTCCACCACTGCGCCACAGCCGCTCG 2753
QY 721 GAGATCAAACTCAAGACACCCCAACTGATGCGCCAACTCAATACCAAGGACCAAGC 780
Db |||||
2754 GAGATCAAACTCAAGACACCCCAACTGATGCGCCAACTCAATACCAAGGACCAAGC 2813
QY 781 AGTGATGAGGAGCTCGCAACATCCGCTCAGGTCGGAGAACGAGAACCCACACACA 840
Db |||||
2814 AGTGATGAGGAGCTCGCAACATCCGCTCAGGTCGGAGAACGAGAACCCACACACA 2873
QY 841 ACTTCTGATGCGGTACCAAGCAAGGCTTTTATCAATCAATGCGACCCCTCCACACA 900
Db |||||
2874 ACTTCTGATGCGGTACCAAGCAAGGCTTTTATCAATCAATGCGACCCCTCCACACA 2933
QY 901 CAACCAAGCAGCCACAGCAAGGAGAAACACACAAACCAATTTCCCAAGATGCTGTGACT 960
Db |||||
2934 CAACCAAGCAGCCACAGCAAGGAGAAACACACAAACCAATTTCCCAAGATGCTGTGACT 2993
QY 961 GAACTAGACAAATAAACAACACTGCAACCGTCCATGCGCCCTCATAAACAACCTACACACA 1020
Db |||||
2994 GAACTAGACAAATAAACAACACTGCAACCGTCCATGCGCCCTCATAAACAACCTACACACA 3053
QY 1021 ATCTCTACTAACAACACTCCAAACACAACTTTCAGGACTCTCTCTGACCAATTAACAAAC 1080
Db |||||
3054 ATCTCTACTAACAACACTCCAAACACAACTTTCAGGACTCTCTCTGACCAATTAACAAAC 3113
QY 1081 ACCACCAATGACACACACAGAGCAATCACTGAAATGAGCAACAGTGCCTCCCTCG 1140
Db |||||
3114 ACCACCAATGACACACACAGAGCAATCACTGAAATGAGCAACAGTGCCTCCCTCG 3173
QY 1141 ATAAACAACCTCGCTCCAAACGGGAATCCACACAGCAAGAGAGCAGCAGCAAAAAA 1200
Db |||||
3174 ATAAACAACCTCGCTCCAAACGGGAATCCACACAGCAAGAGAGCAGCAGCAAAAAA 3233
QY 1201 GSCCGCGGCAACAGCGCACCAACAGCAATGAGCAATTCACAGTCTCTCCCGCCACC 1260
Db |||||
3234 GSCCGCGGCAACAGCGCACCAACAGCAATGAGCAATTCACAGTCTCTCCCGCCACC 3293
QY 1261 CCAGCTGAGTGCACACATCTGTATATTTTCAGAGAAAGCGA 1305
Db |||||
3294 CCAGCTGAGTGCACACATCTGTATATTTTCAGAGAAAGCGA 3338
```

RESULT 5

```
US-10-491-121-32
; Sequence 32, Application US/10491121
; Publication NO. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct padapt Marburg GP (dTM)
```

US-10-491-121-32

```
Query Match 70.8%; Score 1303.4; DB 20; Length 8256;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGACCAACATCTTCTTATCAGTCTTATCTTAACTTCAAGGACAAATAATCTCCCC 60
Db |||||
1431 ATGAGACCAACATCTTCTTATCAGTCTTATCTTAACTTCAAGGACAAATAATCTCCCC 1490
QY 61 ATTTTAGAGATAGTAGTAATAAACAACCCCAAAATGTGGATTGGATTGCTATGCTCCGGAAT 120
Db |||||
1491 ATTTTAGAGATAGTAGTAATAAACAACCCCAAAATGTGGATTGGATTGCTATGCTCCGGAAT 1550
QY 121 CTCAGAGACAGAGAGAGCTCCATCTGTATGGGATTTCAACATGAGTGGGCAAAAAGTTGCT 180
Db |||||
1551 CTCAGAGACAGAGAGAGCTCCATCTGTATGGGATTTCAACATGAGTGGGCAAAAAGTTGCT 1610
QY 181 GATTCCCTTTTGGAGGATCCAAAGGATGGCTTTTCAAGCAGGTGTACTTCCCAAGAAAT 240
Db |||||
1611 GATTCCCTTTTGGAGGATCCAAAGGATGGCTTTTCAAGCAGGTGTACTTCCCAAGAAAT 1670
QY 241 GTTGAGTACACAGAGGGGAGAGAGCCAAACATGCTACAATAATAAGTGTAAACGGATCCC 300
Db |||||
1671 GTTGAGTACACAGAGGGGAGAGAGCCAAACATGCTACAATAATAAGTGTAAACGGATCCC 1730
QY 301 TCTGGAAAAATCCTTGTCTTATAGATCTCTTACCAACATCCGTGACTATCCGAAATGCAAA 360
Db |||||
1731 TCTGGAAAAATCCTTGTCTTATAGATCTCTTACCAACATCCGTGACTATCCGAAATGCAAA 1790
QY 361 ACTATCCATCATATTTCAAGTCAAAACCTCATGCAAGGGATGCGCTTCATTTATG 420
Db |||||
1791 ACTATCCATCATATTTCAAGTCAAAACCTCATGCAAGGGATGCGCTTCATTTATG 1850
QY 421 GGAGCATTTTTCTGTATGATGCGCATTTGCCCTTCAACAATGTACCGAGGCAAAAGTCTTC 480
Db |||||
1851 GGAGCATTTTTCTGTATGATGCGCATTTGCCCTTCAACAATGTACCGAGGCAAAAGTCTTC 1910
QY 481 ACTGAAAGGAAACATAGCAGCTATGATTTCAATAGACAGTGCACAAATGATTTTCG 540
Db |||||
1911 ACTGAAAGGAAACATAGCAGCTATGATTTCAATAGACAGTGCACAAATGATTTTCG 1970
QY 541 CGGCAAGGACAGAGGTACCGTCAATGATCTGACTTCTACTAATAATAATTTGGACAAGT 600
Db |||||
1971 CGGCAAGGACAGAGGTACCGTCAATGATCTGACTTCTACTAATAATAATTTGGACAAGT 2030
QY 601 AGTAAACGGAACGCAAAACGATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAATTC 660
Db |||||
2031 AGTAAACGGAACGCAAAACGATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAATTC 2090
QY 661 ACAAAGAACCAAAACATGTGCTCCGTCCAAATAATCTCCACCACTGCCCCACAGCCGCTCG 720
Db |||||
2091 ACAAAGAACCAAAACATGTGCTCCGTCCAAATAATCTCCACCACTGCCCCACAGCCGCTCG 2150
QY 721 GAGATCAAACTCAACAGCACCCCACTGATGCGCAACCAATCAATACACCGACCAAGC 780
Db |||||
2151 GAGATCAAACTCAACAGCACCCCACTGATGCGCAACCAATCAATACACCGACCAAGC 2210
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCACACA 840
Db |||||
2211 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCACACA 2270
QY 841 ACTTCTGATGCGGTACCAAGCAAGGCTTTTCAACAATGCGACCAACCTCCCAAGATGCTGTGACT 900
Db |||||
2271 ACTTCTGATGCGGTACCAAGCAAGGCTTTTCAACAATGCGACCAACCTCCCAAGATGCTGTGACT 2330
QY 901 CAACCAAGCAGCCACAGAGAGGAGAAACAAACAAACCAATGCGACCCCTCCCTCACCA 960
Db |||||
2331 CAACCAAGCAGCCACAGAGAGGAGAAACAAACAAACCAATGCGACCCCTCCCTCACCA 2390
QY 961 GAACTAGACAAATAAACAACCTGCAACCGTCCATGCGCCCTCATAAACAACCTACCAACA 1020
Db |||||
2391 GAACTAGACAAATAAACAACCTGCAACCGTCCATGCGCCCTCATAAACAACCTACCAACA 2450
```

```
QY 1021 ATCTCTACTAAACAACACCTCCAAAACAACCTTCAGCACTCTCTCTGACCAATTACAAAAC 1080
Db 2451 ATCTCTACTAAACAACACCTCCAAAACAACCTTCAGCACTCTCTCTGACCAATTACAAAAC 2510
QY 1081 ACCACCAATGACAACACACAGAGCACAATCACTGAAATGAGCAACACGAGTGCCTCCCTCG 1140
Db 2511 ACCACCAATGACAACACACAGAGCACAATCACTGAAATGAGCAACACGAGTGCCTCCCTCG 2570
QY 1141 ATAACAACCTCTCTCAACGGGAAATCCACACAGCAAGAGAGCACCAGAGCAAAAAA 1200
Db 2571 ATAACAACCTCTCTCAACGGGAAATCCACACAGCAAGAGAGCACCAGAGCAAAAAA 2630
QY 1201 GGCCTCCGACAAACGGCACAACAGCAAAATGAGCAATTCACAGTCTCTCCCTCCACC 1260
Db 2631 GGCCTCCGACAAACGGCACAACAGCAAAATGAGCAATTCACAGTCTCTCCCTCCACC 2690
QY 1261 CCAGCTCGACTGCACAACATCTGTATATTTTCAGAGAAAGCGA 1305
Db 2691 CCAGCTCGACTGCACAACATCTGTATATTTTCAGAGAAAGCGA 2735

RESULT 6
US-10-267-322-1
; Sequence 1, Application US/10267322
; Publication No. US20030152590A1
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143D
; CURRENT APPLICATION NUMBER: US/10/267,322
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 09/336,910, and US 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
; ORGANISM: Marburg Virus
US-10-267-322-1

Query Match 70.8%; Score 1303.4; DB 16; Length 11460;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTTCTTATCAGTCTTATTAATTCAGGGACAAAAATCTCCCC 60
Db 5940 ATGAAGACCAATGTTTCTTATCAGTCTTATTAATTCAGGGACAAAAATCTCCCC 5999
QY 61 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 6000 ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 6059
QY 121 CTCAGAGAAGACAGACGTCATCTGATGGGATTCACACTAGTGGGGCAAAAAGTTGCT 180
Db 6060 CTCAGAGAAGACAGACGTCATCTGATGGGATTCACACTAGTGGGGCAAAAAGTTGCT 6119
QY 181 GATTCCTTTTGAGGCATCCAGCGATGGCTTTTCAGGACAGGTGTACTCCCAAGAT 240
Db 6120 GATTCCTTTTGAGGCATCCAGCGATGGCTTTTCAGGACAGGTGTACTCCCAAGAT 6179
QY 241 GTTGAGTACACAGAGGGGAGGAAGCAAAAACATGCTACAATATAAGTGAACGGATCCC 300
Db 6180 GTTGAGTACACAGAGGGGAGGAAGCAAAAACATGCTACAATATAAGTGAACGGATCCC 6239
QY 301 TCTGGAATAATCCTTGCTGTAGATCTCTCTCAACATCCGATGATTCGGAATTCGAAA 360
Db 301 TCTGGAATAATCCTTGCTGTAGATCTCTCTCAACATCCGATGATTCGGAATTCGAAA 360
```

```
Db 6240 TCTGGAATAATCCTTGCTGTAGATCTCTCTCAACATCCGATGATTCCTAAATGCAAA 6299
QY 361 ACTATCCATCATATTTCAAGGTCAAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
Db 6300 ACTATCCATCATATTTCAAGGTCAAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 6359
QY 421 GGAGCATTTTTTCTGTATGATCGCATTCCTCCACAACATGTATCCGAGGCAAAAGTCTTC 480
Db 6360 GGAGCATTTTTTCTGTATGATCGCATTCCTCCACAACATGTATCCGAGGCAAAAGTCTTC 6419
QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCTCG 540
Db 6420 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCTCG 6479
QY 541 CGGCAAGGACAGAGGTACCGTCATATGAATCTGACTTCTTACTATAAATAATTTGGAACAAGT 600
Db 6480 CGGCAAGGACAGAGGTACCGTCATATGAATCTGACTTCTTACTATAAATAATTTGGAACAAGT 6539
QY 601 AGTAAACGGAAACGCAAAACGAATGACACTGGATGTTTGGGGCTCTTTCAAGAATATCAAAATCT 660
Db 6540 AGTAAACGGAAACGCAAAACGAATGACACTGGATGTTTGGGGCTCTTTCAAGAATATCAAAATCT 6599
QY 661 ACAAGAACCAAAACATGTGCTCCGTCCAAAATACCTCCACCACTGCCCCACAGCCCGTCCG 720
Db 6600 ACAAGAACCAAAACATGTGCTCCGTCCAAAATACCTCCACCACTGCCCCACAGCCCGTCCG 6659
QY 721 GAGATCAAACTCACAAAGCAGCCCAACTGATGCCACCAAACTCAATACCCAGGACCCCAAGC 780
Db 6660 GAGATCAAACTCACAAAGCAGCCCAACTGATGCCACCAAACTCAATACCCAGGACCCCAAGC 6719
QY 781 AGTGATGATGAGGACTCTCGAAACATCGGGCTCAGGGTCGGAGAAACGAGAACCCCAACA 840
Db 6720 AGTGATGATGAGGACTCTCGAAACATCGGGCTCAGGGTCGGAGAAACGAGAACCCCAACA 6779
QY 841 ACTTCTGATGCGGTCAACCAAGCAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACC 900
Db 6780 ACTTCTGATGCGGTCAACCAAGCAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACC 6839
QY 901 CAACCAAGCAGCCACAGAGAGGAGAAACAACCAAACTTCCCAAGATGCTGTGACT 960
Db 6840 CAACCAAGCAGCCACAGAGAGGAGAAACAACCAAACTTCCCAAGATGCTGTGACT 6899
QY 961 GAACTAGACAAAAATAACAACATGACAAACCGTCCATGCCCTCATAAACAATAACCA 1020
Db 6900 GAACTAGACAAAAATAACAACATGACAAACCGTCCATGCCCTCATAAACAATAACCA 6959
QY 1021 ATCTCTACTAACACACCTCCAAAACAACATTCAGCACTCTCTGACCAATTCACAAAC 1080
Db 6960 ATCTCTACTAACACACCTCCAAAACAACATTCAGCACTCTCTGACCAATTCACAAAC 7019
QY 1081 ACCACCAATGACAACACACAGAGCACAATCACTGAAAAATGAGCAAAACAGTGCCTCCCTCG 1140
Db 7020 ACCACCAATGACAACACACAGAGCACAATCACTGAAAAATGAGCAAAACAGTGCCTCCCTCG 7079
QY 1141 ATAACAACCTCTCTCAACGGGAAATCCACACAGCAAGAGAGCACCAGAGCAAAAAA 1200
Db 7080 ATAACAACCTCTCTCAACGGGAAATCCACACAGCAAGAGAGCACCAGAGCAAAAAA 7139
QY 1201 GGCCTCCGACAAACGGCACAACAGCAAAATGAGCAATTCACCAAGTCTCTCCCTCCACC 1260
Db 7140 GGCCTCCGACAAACGGCACAACAGCAAAATGAGCAATTCACCAAGTCTCTCCCTCCACC 7199
QY 1261 CCAGCTCGACTGCACAACATCTGTATATTTTCAGAGAAAGCGA 1305
Db 7200 CCAGCTCGACTGCACAACATCTGTATATTTTCAGAGAAAGCGA 7244
```

```
RESULT 7
US-10-066-506A-9
; Sequence 9, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
```

APPLICANT: Hevey, Michael C.
APPLICANT: Schmaljohn, Alan, L.
TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
FILE REFERENCE: 003/243/SAP
CURRENT APPLICATION NUMBER: US/10/066,506A
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/267,522
PRIOR FILING DATE: 2001-01-31
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 9
LENGTH: 2051
TYPE: DNA
ORGANISM: Marburg virus strain Musoke
FEATURES:
OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 and Marburg virus Glycoprotein 2
OTHER INFORMATION: Glycoprotein 2
US-10-066-506A-9

Query Match 69.2%; Score 1274; DB 15; Length 2051;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1296; Conservative 0; Mismatches 0; Indels 12; Gaps 1;
QY 1 ATGAAGACCATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACACAAAATCTCCCC 60
DB 10 ATGAAGACCATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACACAAAATCTCCCC 69
QY 61 ATTTTAGAGTAGTAGTAATAATCAACCCCAAAATGTGGATGTCTCCGGAAT 120
DB 70 ATTTTAGAGTAGTAGTAATAATCAACCCCAAAATGTGGATGTCTCCGGAAT 129
QY 121 CTCGAGACAGAGACGTCCATCTGATGGGATTCACCTGAGTGGGCAAAAGTTGCT 180
DB 130 CTCGAGACAGAGACGTCCATCTGATGGGATTCACCTGAGTGGGCAAAAGTTGCT 189
QY 181 GATTCCTTTTGAGGATCCCAAGGATGGGCTTTTCAGGACAGGTGATCTCCCAAGAT 240
DB 190 GATTCCTTTTGAGGATCCCAAGGATGGGCTTTTCAGGACAGGTGATCTCCCAAGAT 249
QY 241 GTTGTAGTACAGAGGGGAGGAGCAAAACATGCTACAATATAAGTGTAAACGGATCCC 300
DB 250 GTTGTAGTACAGAGGGGAGGAGCAAAACATGCTACAATATAAGTGTAAACGGATCCC 309
QY 301 TCTGGAATATCTTGTGTAGATCTCTCAACATCCGTGACTATCCGAAATGCAAA 360
DB 310 TCTGGAATATCTTGTGTAGATCTCTCAACATCCGTGACTATCCGAAATGCAAA 369
QY 361 ACTATCCATCATTTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
DB 370 ACTATCCATCATTTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 429
QY 421 GGAGCATTTTCTGTATGATGGCATTTGCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 480
DB 430 GGAGCATTTTCTGTATGATGGCATTTT-----ATGTACCGAGGCAAAAGTCTTC 477
QY 481 ACTGAAGGAAACATAGCAGTATGATGTCAATAGACAGTGCACAAAATGATTTTCG 540
DB 478 ACTGAAGGAAACATAGCAGTATGATGTCAATAGACAGTGCACAAAATGATTTTCG 537
QY 541 CGGACAGGACNAGGTTACCGTCAATGATCTGACTTCTACTAATAATATGGCAAGT 600
DB 538 CGGACAGGACNAGGTTACCGTCAATGATCTGACTTCTACTAATAATATGGCAAGT 597
QY 601 AGTAACGGAAACGAAACGATGATGTTTCGGGCTCTTCAAGAAATACAAATCT 660
DB 598 AGTAACGGAAACGAAACGATGATGTTTCGGGCTCTTCAAGAAATACAAATCT 657
QY 661 ACAAGAACCAAAACATGTGCTCGGTCCAAATACTCCACCACTGCCACACCCGCTCG 720
DB 658 ACAAGAACCAAAACATGTGCTCGGTCCAAATACTCCACCACTGCCACACCCGCTCG 717
QY 721 GAGATCAAACTCAAGACACCCCACTGATGACCAAACTCAATACAGGACCCCAAGC 780

DB 718 GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCATATACCGGACCCCAAGC 777
QY 781 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACACA 840
DB 778 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACACA 837
QY 841 ACTTCTGATGCGGTCCCAAGCAAGGGCTTTCATCAAAATGCCACACCTCCCTCACCAC 900
DB 838 ACTTCTGATGCGGTCCCAAGCAAGGGCTTTCATCAAAATGCCACACCTCCCTCACCAC 897
QY 901 CAACCAAGCAGCCACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
DB 898 CAACCAAGCAGCCACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 957
QY 961 GAATAGACAAAATAACACAACTGCAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 958 GAATAGACAAAATAACACAACTGCAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
QY 1021 ATCTCTACTTAACAACACCTCCAAACACAACTTTCAGCACTCTCTCTGCAACATTAACAAAC 1080
DB 1018 ATCTCTACTTAACAACACCTCCAAACACAACTTTCAGCACTCTCTCTGCAACATTAACAAAC 1077
QY 1081 ACCACCAATGACACACACAGAGCACAATCACTGAAAATGAGCAAAACCAAGTCCCTCTCG 1140
DB 1078 ACCACCAATGACACACACAGAGCACAATCACTGAAAATGAGCAAAACCAAGTCCCTCTCG 1137
QY 1141 ATAAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGACACAGCAGCAAAAAA 1200
DB 1138 ATAAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGACACAGCAGCAAAAAA 1197
QY 1201 GGCCCCGCCCAACCGGACCAACACAGCAAAATGAGCAATTCACCACTCTCTCCCTCCACCC 1260
DB 1198 GGCCCCGCCCAACCGGACCAACACAGCAAAATGAGCAATTCACCACTCTCTCCCTCCACCC 1257
QY 1261 CCCAGCTCGACTGCAACCAACATCTTGTATATTTTCAGAAAGAGCGATCG 1308
DB 1258 CCCAGCTCGACTGCAACCAACATCTTGTATATTTTCAGAAAGAGCGATCG 1305

RESULT 8

US-10-811-353-2
Sequence 2, Application US/10811353
Publication No. US20050112098A1
GENERAL INFORMATION:
APPLICANT: McCray, Paul B.
APPLICANT: Sanders, David A.
APPLICANT: Jeffers, Scott A.
APPLICANT: Davidson, Beverly L.
APPLICANT: Simm, Patrick L.
TITLE OF INVENTION: PSEUDOTYPED VIRUSES AND METHODS FOR THEIR USE
FILE REFERENCE: 290.00670120
CURRENT APPLICATION NUMBER: US/10/811,353
CURRENT FILING DATE: 2004-03-26
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 2046
TYPE: DNA
ORGANISM: Marburg virus glycoprotein
US-10-811-353-2

Query Match 62.4%; Score 1148.2; DB 21; Length 2046;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
QY 1 ATGAAGACCATGTTTCTTATCAGTCTTATCTTAATTTCAAGGACACAAAATCTCCCC 60
DB 1 ATGAAGACCATGTTTCTTATCAGTCTTATCTTAATTTCAAGGAGTAAAAAATCTCCCT 60
QY 61 ATTTTAGAGTAGTAGTAATAATCAACCCCAAAATGGAATTCGGTATGTCGCGAACT 120
DB 61 ATTTTAGAGTAGTAGTAATAATCAACCCCAAAATGGAATTCGGTATGTCGCGAACT 120

121 CTCGAGACAGAGACGTCCATCTGATGGGATTCACTGAGTGGGCAAAAAGTTGCT 180
121 CTCGAGACAGAGATGCTCATCTGATGGGATTCACTGAGTGGGCAAAAAGTTGCT 180
181 GATTCCTCTTGGAGCATCCAAAGGATGGGCTTTTCAGGACAGGTGATCCTCCCAAGAT 240
181 GATTCCTCTTGGAGCATCCAAAGGATGGGCTTTTCAGGACAGGTGATCCTCCCAAGAT 240
241 GTTGAGTACACAGAGGGGAGGAAGCAAAAATGCTACAATATAGTGTAAACGGATCCC 300
241 GTTGAGTATACAGAGGGGAGGAGCAAAAATGCTACAATATAGTGTAAACGGATCCC 300
301 TCTGMAAATCTTCTGTTAGATCTCTCAACATCCGTGACTATCCGAAATGCAAA 360
301 TCTGMAAATCTTCTGTTAGATCTCTCAACATCCGTGACTATCCGAAATGCAAA 360
361 ACTATCCATCATATCAAGGTCAAAACCTCATGACAGGGATCGCCCTCATTTATGG 420
361 ACTATCCATCATATCAAGGTCAAAACCTCATGACAGGGATCGCCCTCATTTATGG 420
421 GGAGCATTTTCTGTATGATCGCATTTGCTCCAAACAATGTACGGAGCAAAAGTCTTC 480
421 GGAGCATTTTCTGTATGATCGCATTTGCTCCAAACAATGTACGGAGCAAAAGTCTTC 480
481 ACTGAAGGAAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTTCTCG 540
481 ACTGAAGGAAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTTCTCG 540
541 CGGCAAGGACAGGGTACCGTACATGATCTGACTTCTACTAATAATATTTGGACAGT 600
541 AGGCAAGGACAGGGTACCGTACATGATCTGACTTCTACTAATAATATTTGGACAGT 600
601 AGTAAACGGAACGAAACGATGATGATGTTTGGGCTCTTCAAGAAATACAAATCT 660
601 AACAATGGAACACAAACGATGATGATGTTTGGGCTCTTCAAGAAATACAAATCT 660
661 ACAAGAAACAAACATGCTCGTCCAAATACCTTCAACATGCTGACACAGCCCGTCCG 720
661 ACGAAGATCAAAACATGCTCGTCCAAATACCTTCAACATGCTGACACAGCCCGTCCG 720
721 GAGATCAACTCACAGACCCCACTGATGATGATGATGATGATGATGATGATGATGATGAT 780
721 GAGATCAAAACCAAGACCCCACTGATGATGATGATGATGATGATGATGATGATGATGAT 780
781 AGTATGATGAGGACCTCTCAACATCCGCTCAGGTCGGGAGACGAAACCCACACA 840
781 AATGATGATGAGGACCTCTCAACATCCGCTCAGGTCGGGAGACGAAACCCCTATACA 840
841 ACTTCTGATGCGGTACCAAGCAAGGCTTTTCAACAAATGCGCACCACTCCCTCACCA 900
841 ACTTCTGATGCGGTACCAAGCAAGGCTTTTCAACAAATGCGCACCACTCCCTCACCA 900
901 CAACCAAGCACCCACAGCAAGGAGGAACACACAAACCATTTCCCAAGTCTGTGACT 960
901 CAACCAAGCACCCACAGCAAGGAGGAACACACAAACCATTTCCCAAGTCTGTGACT 960
961 GAACTAGACAAAAATAACAACTGCAACCGTCCATGCCCCCTCATAACTACTACCACA 1020
961 GAAACCAACAAACCAACCAAGGAGGAACACACAAACCATTTCCCAAGTCTGTGACT 1020
1021 ATCTCTACTAAACAACTCTCAACAACTCTTCAACAACTCTTCTGACCACTTACAAAC 1080
1021 ATCTCTACTAAACAAACCTCTCAACAACTCTTCAACAACTCTTCTGACCACTTACAAAC 1080
1081 ACCCAACATGACACACAGGACCACTACTGAAATGACAAACCATTTCCCAAGTCTGTGACT 1140
1081 ACCCAACATGACACACAGGACCACTACTGAAATGACAAACCATTTCCCAAGTCTGTGACT 1140
1141 ATAACAAACCTCTCTCAACAGGAAATCCACACAGCAAGGAGCAACCAAGGAGCAAAAA 1200
1141 AAAACAAACCTCTCTCAACAGGAAATCTTCAACAGCAAGGAGCACTTCAACAGCAAGAA 1200
1201 GGGCCCGCCCAACAGGCAACAAACAGCAAAATAGGACATTTTCAACAGTCTCTCCCCCACC 1260

Db 1201 GGGCCCGCCCAACAGGCAACAAATAGGACATTTTAAACAGTCTCTCCCCCACC 1260
Qy 1261 CCCAGCTCGATCGACCAACATCTTGTATATTTTTCAGAAAGAAAGCGA 1305
Db 1261 CCCAACCCGACCAACACATCTTGTATATTTTCAGAAAGAAAGCGA 1305
RESULT 9
US-10-353-856-27
; Sequence 27, Application US/10353856
; Publication No. US20030215794A1
; GENERAL INFORMATION:
; APPLICANT: Kawaoka, Yoshihiro
; APPLICANT: Jasenosky, Luke D.
; APPLICANT: Neumann, Gabriele
; APPLICANT: Wisconsin Alumni Research Foundation
; TITLE OF INVENTION: Filovirus Vectors and No. US20030215794A1infectious Filovirus-Base
; FILE REFERENCE: 800.032US1
; CURRENT APPLICATION NUMBER: US/10/353,856
; PRIOR FILING DATE: 2003-01-29
; PRIOR FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 19112
; TYPE: DNA
; ORGANISM: Marburg virus
US-10-353-856-27
Query Match 62.4%; Score 1148.2; DB 17; Length 19112;
Best Local Similarity 92.5%; Pred. No. 0;
Matches 1207; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
Qy 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGACAAAAATCTCCCC 60
Db 5940 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGACAAAAATCTCCCC 5999
Qy 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGGAATTCGGATTCGTCCGGAAT 120
Db 6000 ATTTTAGAGATAGTAGTAACATCAACCCCAAAATGGAATTCGGATTCGTCCGGAAT 6059
Qy 121 CTCGAGACAGAGACGTCCATCTGTATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
Db 6060 CTCGAGACAGAGAGATGTCATCTGTATGGGATTCACATGAGTGGGCAAAAAGTTGCT 6119
Qy 181 GATTCCTCTTGGAGGATCCAAAGGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db 6120 GATTCCTCTTGGAGGATCCAAAGGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 6179
Qy 241 GTTGAGTACACAGAGGGGAGGAAGCAAAAATGCTACAATATAGTGTAAACGGATCCC 300
Db 6180 GTTGAGTATACAGAGGGGAGGAGCAAAAATGCTACAATATAGTGTAAACGGATCCC 6239
Qy 301 TCTGMAAATCTTCTGTTAGATCTCTCAACATCCGTGACTATCCGAAATGCAAA 360
Db 6240 TCTGMAAATCTTCTGTTAGATCTCTCAACATCCGTGACTATCCGAAATGCAAA 6299
Qy 361 ACTATCCATCATATTTCAAGGTCAAAAATGCTATGACAGGGATCGCCCTCATTTATGG 420
Db 6300 ACTATCCATCATATTTCAAGGTCAAAAATGCTATGACAGGGATCGCCCTCATTTATGG 6359
Qy 421 GGAGCATTTTCTGTATGATCGCATTTGCTCCAAACAATGTTACCGAGGCAAAAGTCTTC 480
Db 6360 GGAGCATTTTCTGTATGATCGCATTTGCTCCAAACAATGTTACCGAGGCAAAAGTCTTC 6419
Qy 481 ACTGAAGGAAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTTCTCG 540
Db 6420 ACTGAAGGAAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTTCTCG 6479
Qy 541 CGGCAAGGACAGGGTACCGTACATGATCTGACTTCTACTAATAATATTTGGACAGT 600

Db 6480 AGGCAAGGACAGGGGTACCGTACATGAATCTGACTTCTACTATAATAATATTGGACAGT 6539
Qy 601 AGTAACGGAACCAACGAATGACACTGATGTTTCGGGGCTCTTCAAGAAATACAAATCTT 660
Db 6540 AACAAATGGAAACAAACGAATGACACTGATGTTTCGGGTCTCTTCAAGAAATACAACTCC 6599
Qy 661 ACAGAAACCAAAATGCTCGCTCCAAATACCTCCACCACTGCGCCACAGCCCGTCCG 720
Db 6600 AGAAGAAATCAAAATGCTCGCTCCAAATACCTCCACCACTGCGCCACAGCCCGTCCA 6659
Qy 721 GAGATCAAACTCAAGCAACCCAACTGATGTCACCAAACTCAATACCAAGCCCAAGC 780
Db 6660 GAGATCAAACTCAAGCAACCCAACTGATGTCACCACTCAATACCAAGCCCAAGC 6719
Qy 781 AGTGATGATGAGGACTCGCAACATCGGCTCAGGGTCGGAGAAAGAGAAACCCACACA 840
Db 6720 AATGATGATGAGGACTCTATAATCCGGTTCAGGGTCGGAGAAACAGGAAACCCCTATACA 6779
Qy 841 ACTTCTGATGCGGTCAACCAAGCAAGGGTCTTATCAACAAATGCCACCCACTCCCTCACCA 900
Db 6780 ACTTCAGATGCGGTCACTAAGCAAGGGTCTTATCAACAAATGCCACCCACTCCCTCACCA 6839
Qy 901 CAACCAAGCAACCCACAGCAAGGAGAAACCAACCAAACTTCCCAAGATGCTGTGACT 960
Db 6840 CAACCAAGCAACCCACAGCAAGGAGAAACCAACCAAGACTTCCCAAGGACTGTGACT 6899
Qy 961 GAACTAGACAAATAACAACTGCAACCGTCCATGCGCCCTCATATAACACTACCAACA 1020
Db 6900 GAAACCAACCAACCAACCAACCGTCCATGCGCCCTCATATAACACTACCAACA 6959
Qy 1021 ATCTCTACTACCAACCTCCAAACCAACCTGACACTCTCTGACCACTTACCAAAAC 1080
Db 6960 ATCTCTACTACCAACCTCCAAACCAACCTGACACTCTCTGATATCACTACAAAC 7019
Qy 1081 ACCCAATATGACACACAGACCAATCACTGAAATATGACAAACCAAGTGCCTCCCTCG 1140
Db 7020 ACCCAATATGACACACAGACCACTGAAATATGACAAACCAAGTGCCTCCCTCG 7079
Qy 1141 ATAAACCTGCTCCAAACGGGAAATCCACCAACAGCAAGAGACCAAGCAAGCAAA 1200
Db 7080 AAAACAACTGCTCCAAACAGGAAATCTTACCAACAGCAAGAGCACTTAAACACAGAA 7139
Qy 1201 GCGCCGCGCACAAACGACCAACAGCAAGCAAGTATGAGCACTTCAACCACTCCCTCC 1260
Db 7140 GCGCCGCGCACAAACGACCAACAGCAAGCAAGTATGAGCACTTAAACCACTCCCTCC 7199
Qy 1261 CCGAGCTGCACTGCAACATCTCTGATATTTTCAAGAAAGCGA 1305
Db 7200 CCCAACCCGACCAACACATCTTGTATATTTTCAAGAAAGCGA 7244

RESULT 10

US-10-066-506A-7
; Sequence 7, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 7
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Raven Glycoprotein

OTHER INFORMATION: Marburg virus strain Musoke Glycoprotein 2
US-10-066-506A-7

Query Match 42.4%; Score 780; DB 15; Length 2046;
Best Local Similarity 74.8%; Pred. No. 3e-232;
Matches 978; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

Qy 1 ATGAAGACCAATGCTTCTTATCAGTCTTATCTTAATTCAGGACCAAAATCTCCCC 60
Db 1 ATGAAGACCAATATATTTCTGATTTAGTCTCAATTTTAAATCCAAAGTATAAAATCTCCCT 60
Qy 61 ATTTTACAGATAGCTAGTAATAATCAACCCCAAAATGGAATTCGGTATGCTCCGGAAT 120
Db 61 GTTTTAGAATTTGCTAGTAACAGCAACCTCAAGATGTAGATTGAGTGTCTCCGGAACC 120
Qy 121 CTCAGAGACAGAAAGCGTCCATCTGATGGGATTCACATGAGTGGGCAAAAGATTGCT 180
Db 121 CTCAGAGACAGAAAGATGTTTCTGATGGGATTTTACATGAGTGGGCAAAAGATTGCT 180
Qy 181 GATTCCCTTTGAGGAGATCCAGGATGGCTTTTACAGACAGGTGTACCTCCCAAGAT 240
Db 181 GATTCCCTTTGAGGAGATCTAAACGATGGCTTTTACAGACAGGTGTTCCTCCCAAGAAC 240
Qy 241 GTTCAGTACACAGAGGGGAGAGCCAAACATGCTTACATATTAAGTGTAAACGATCC 300
Db 241 GTTCAGTACACAGAGGGGAGAGCCAAACATGCTTACATATTAAGTGTAAACGATCC 300
Qy 301 TCTGGAAATCTCTTCTGCTTGTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Db 301 TCTGGAAATCTCTTCTGCTTGTAGATCTCTTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Qy 361 ACTATCATATATTAAGGTCAAAACCTCATGACAGGGGATCGGCTTCATTTATGG 420
Db 361 ACTGTTTATCATATTAAGGTCAAAACCTCATGACAGGGGATTCGCTTCATTTATGG 420
Qy 421 GGAGCATTTTCTGATGATGCGTCTTACCAACATGTTACCGAGGCAAGGTCTTC 480
Db 421 GGAGCATTTTCTGATGATGCGTCTTACCAACATGTTACCGAGGCAAGGTCTTC 480
Qy 481 ACTGAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTCTCG 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTCT 540
Qy 541 CGGCAAGGACAGGGTACCGTCAATATGAACTGCTTCTACTATAAATATTTGGACAAGT 600
Db 541 AGGCAAGGACAGGGTATTCGTCACTGAACTTGACCTCCACCAATAAATATTTGGACAAGC 600
Qy 601 AGTAACGGAACGCAACGAAATGACATGGAATGTTTCGGGCTCTTCAAGATACAAATCT 660
Db 601 AGCAATGAACGCAAGAAATGATACGGGATGTTTGGCATTCCTCCAAAGAAATACAACTCC 660
Qy 661 ACAAGGAAACCAACATGCTGCTCCGTCCTCAAAATACCTCCACCTGCGCCACAGCCGTCGG 720
Db 661 ACAAGCAATCAACATGCGCTCCGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACTCCG 720
Qy 721 GAGATCAAACTCAAGCAACCCCAACTGATGTCACCAAACTCAATACCAAGCAAGCAAGC 780
Db 721 AGCATTCACCTCAAAATACCTCAAAATTAATGCTGCTAAAATCTGGAACCTATGAAACCAAGT 780
Qy 781 AGTGATGATGAGGACTCGCAACATCCGGCTCAGGGTCGGGAGAACAGAAACCCCAACA 840
Db 781 AGCGACGATGAGGACCTTATGATTTTCGGGCTCAGGATCTGGAGAAACAGGGGCCCCACACA 840
Qy 841 ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTTATCAACAAATGCCACCCACTCCCTCACCA 900
Db 841 ACTCTTAATGATGCTCACTGAAACAAACATCTGTCACAAATATTGTCACCTCTCTCACTA 900
Qy 901 CAACCAAGCAAGCCACAGCAAGGAGGAAACAAACCAACCTTCCCAAGATGCTGTGACT 960
Db 901 CATCAAGCACTCAACACATGAGCAAAACAGATACGAATCTCTCCGACATGCTGTAAT 960
Qy 961 GAACTAGCAAAATAACACAACTGCAACACCGTCCATGCGCCCTCATATAACACTACCAACA 1020

Db	961	GAGCACAATGGACCGGACCCAAACAACACACCGACGACGGTCTCTCAACATACTAATACA	1020
Qy	1021	ATCTCTATAACACACACCTCCAAACACAACTTTCAGCACTCTCTCTGACCAATTAACAAAC	1080
Db	1021	ACTCCACCTTATAACACTCTCAAGTACAACCTCTAGTACTCTTCTCCCTCTCCAACCCGCAAC	1080
Qy	1081	ACACCAATGACAAACACACAGACGACAATCACTGMAAATGAGCAGAAACAGTGGCCCCCTCG	1140
Db	1081	ATACCAAAATATGATAACAACGTGAATCTAGCAGAAAGCGAAACAACCAATGCTCAGTTG	1140
Qy	1141	ATAACAACCTTCCTCCAAACGGGAAATCCCAACACACAGCAAAAGACCAAGCAGCAAAAAA	1200
Db	1141	AACACAACCTCTAGATCCAAACAGAAATCCCAACACAGCAAGACACCAACAGCACAACC	1200
Qy	1201	GGCCCCGACACACGGGACCAACACACAGCAAAATGAGGATTTACCAGTCTCTCCCCCACC	1260
Db	1201	AACATCATCATGACGACATCAGATATATACAGCAAAACACCCCAAAATTTCTTCGGAT	1260
Qy	1261	CCAGCTCGACTGCACAACACTCTGTATATTTTTCAGAAAGAACGCATCG	1308
Db	1261	TCTAGTCGACACCCGCCCTCTCTATATACTTTAGAAAGAACGATCG	1308

RESULT 11

US-10-066-506A-13

; Sequence 13, Application US/10066506A

; Publication No. US20030108560A1

; GENERAL INFORMATION:

; APPLICANT: Grogan, Case C.

; APPLICANT: Hevey, Michael C.

; APPLICANT: Schmaljohn, Alan, L.

; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein

; FILE REFERENCE: 003/243/SAP

; CURRENT APPLICATION NUMBER: US/10/066,506A

; CURRENT FILING DATE: 2002-01-31

; PRIOR APPLICATION NUMBER: 60/267,522

; PRIOR FILING DATE: 2001-01-31

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Apple Macintosh Microsoft Word 6.0

; SEQ ID NO 13

; LENGTH: 2046

; TYPE: DNA

; ORGANISM: Marburg virus strain Raven

; FEATURE:

; OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 a

; OTHER INFORMATION: Glycoprotein 2

US-10-066-506A-13

Query Match	42.4%	Score 780;	DB 15;	Length 2046;
Best Local Similarity	74.8%;	Pred. No. 3e-232;		
Matches 978;	Conservative 0;	Mismatches 330;	Indels 0;	Gaps 0;
Qy	1	ATGAAGACCACATGTTTCCITATCAGTCTTATCTTAATTCAGGGACAAAATACTCCCC	60	
Db	1	ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGATATAAAACTCTCCCT	60	
Qy	61	ATTTTACAGATAGCTAGTAAATAATCAACCCCAAATGTGGATTCCGTATGTCCTCGGAAC	120	
Db	61	GTTTTAGAATTTGTAGTAAACAGCAACCTCAAGATGTAGATTCAAGTGTCTCCGGAAC	120	
Qy	121	CTCCAGAAGACAGAAGACGTCCTATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	180	
Db	121	CTCCAAAGACAGAAGATGTTTCATCTCATGGGATTTACACTGAGTGGGCAAAAAGTTGCT	180	
Qy	181	GATTCCTCTTTGGAGGCATCCAAAGCATGGGCTTTCAGGACAGGTGTAACCTCCCAAGAT	240	
Db	181	GATTCCTCTTTGGAGGATCTTAAACGATGGGCTTTCAGGACAGGTGTTCTCTCCCAAGAA	240	
Qy	241	GTTGAGTACACAGAGGGGAGGACCAACATGCTACATATAGTGTAACGGATCCC	300	
Db	241	GTTGAGTATACGGAAGGAGAAAGCCAAAACATGTTACAATTAAGTGTAAACAGACCT	300	
Qy	301	TCTGGAAAATCCTTGCTGTTAGATCTCTTACCACACATCCGCTGACTATCCGAAATGCAAA	360	

D	b		301	TCTGGAAAAATCCCTTGCTGCTGGAGTCTCCCGATAATTCGCGATTACCCCTAAATGTAAA	360
Q	y		361	ACTATCCATCATATTTCAAGGTCAAACCCCTCATGSCACAGGGGATCGCCCTTCATTTATGG	420
D	b		361	ACTGTTTCATCATATTTCAAGGTCAAACCCCTCATGSCACAGGGGATTCGCCCTCCATTTGTGG	420
Q	y		421	GAGACATTTTTCTGTATGATCGATTCGCCTCCACAACANTGTACCGAGCAAGCTTTC	480
D	b		421	GGGGCATTTTTCTGTATGATCGGTTCGCCTTCACAACANTGTACCGAGCAAGGCTTTC	480
Q	y		481	ACTGAAGGGAAATAGCAGCTATGATTTGTAATAAGACAGTGCACAAAATGATTTTCTCG	540
D	b		481	ACTGAAGGAATATAGCAGCTATGATTTGTAATAAGACAGTTCACAGAATGATTTTTTCT	540
Q	y		541	CGGCAAGGACAAGGGTACCGTCAATGAAATCTGACTTCTACTAATAAATTTGGACAAGT	600
D	b		541	AGGCAAGGACAAGGGTATTCGTCAATGAACTTGACCCTCCAACAATAAATTTGGACAAGC	600
Q	y		601	AGTAACGGACGCAAAAGAAATGACACTGGAGTGTTCGGCGCTCTTCAAGAATAACAATPCT	660
D	b		601	AGCAATGAACGCGAGAAATGATA CGGATGTTTTGGCATCTCTCCNAGAAATACNACTCC	660
Q	y		661	ACAAAGAACCAAAATATGTGTCTCGTCCGTCOAANAATACCTCCACACTGGCCACAGCCCCGTCG	720
D	b		661	ACAAACAATCAAAATGTCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCCG	720
Q	y		721	GAGATCAAACTCACNAGCACCCCAACTGATGCCACCANAACTCAATACCGAGCCCAAGC	780
D	b		721	AGCATTCACCTCAANAATACTCAAAATTAATATCTGTAAATCTGGAACATAAGAACCCCAAGT	780
Q	y		781	AGTGATGATGAGGACCTTCGCAACATCGGCTCAGGGTCCGGAGAAAGAGAAACCCCACACA	840
D	b		781	AGCGAGATGAGACCTTATGATTTCCGGCTCAGGATCTGAGAAACAGGGGCCCCACACA	840
Q	y		841	ACTTCTGATCGGTACCAAAGCAAGGGCTTTCATCAACAATGCGCA CCCACTCCCTACCA	900
D	b		841	ACTCTTAATGTAGTCACTGAAACAGAAACAATCGTCAACAATATTGTGTCACTCTCTTCACTA	900
Q	y		901	CMAACAGCAGCCACAGCAGAGGAGGAAACAAACACAAACCATTTCCCAAGATGCTGTCACT	960
D	b		901	CATCCAGACCTTCAACAATAGCAAAACAGTACGNATCTCTTCCGACATGCTGTAACT	960
Q	y		961	GAACTAGACAAAATTAACAACTGCAACACCGTCCATGCCCCCTCATACACTACCA	1020
D	b		961	GAGCACAATGGAAACCGACCCCAACAAACAAACACAGCAACGCTCTCTCAACAATACTAATACA	1020
Q	y		1021	ATCTCTACTAACACACCTCCAAACACAACCTTCAGGACTCTCTCTGCAACCATTTACAAAC	1080
D	b		1021	ACTCCCACTTAACTCTCTCAAGTACAACTCAGTACTCTTCTCCCTCCCAACCCGCAAC	1080
Q	y		1081	ACCACCAATGACAAACACAGAGCACAATCACTGAAAATGAGCAAAACGAGTGCCTCCCTCG	1140
D	b		1081	ATCACCANAATGATACACAACGTGAATAGCAGAAAGCGAACAAACCAATGCTCAGTTG	1140
Q	y		1141	ATAACAAACCTTCGCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCACCAGCAGCAAAA	1200
D	b		1141	AACACAATCTAGATCCAAACAGAAATTTCCCAACACAGGACAAGACACCAACAGCACAACC	1200
Q	y		1201	GGCCCCGCCAACAGGCAACCAACACAGCAAAATGAGCATTTTCAACAGTCTCTCCCCCACC	1260
D	b		1201	AACATCATCATGACGACNTCAGATATAAACAGCAAAACCCCAACAAATTTCTCTCCGAT	1260
Q	y		1261	CCCAGCTCGACTGCACAACATCTTTGTATATTTTCAGAAAGGCGATCG	1308
D	b		1261	TCTAGTCCGACAAACCCGCCCTCTCTATATATCTATCTTTAGAAAGAAAGATCG	1308

RESULT 12
US-10-491-121-42
; Sequence 42, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:

; APPLICANT: NABEL, GARY									
; APPLICANT: YANG, ZHI-YONG									
; APPLICANT: SULLIVAN, NANCY									
; APPLICANT: SANCHEZ, ANTHONY									
; TITLE OF INVENTION: Development of a Preventive Vaccine for									
; TITLE OF INVENTION: Filovirus Infection in Primates									
; FILE REFERENCE: NIH221-001NP									
; CURRENT APPLICATION NUMBER: US/10/491,121									
; PRIOR FILING DATE: 2004-03-26									
; PRIOR APPLICATION NUMBER: PCT/US02/30251									
; PRIOR FILING DATE: 2002-09-24									
; PRIOR APPLICATION NUMBER: US 60/326476									
; PRIOR FILING DATE: 2001-10-01									
; NUMBER OF SEQ ID NOS: 52									
; SOFTWARE: FastSEQ for Windows Version 4.0									
; SEQ ID NO 42									
; LENGTH: 6902									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Synthetic Construct pVR1012x/s Marburg (codon optimized)									
US-10-491-121-42									
Query Match 37.4%; Score 688; DB 20; Length 6902;									
Best Local Similarity 70.5%; Pred. No. 3.7e-203;									
Matches 919; Conservative 0; Mismatches 385; Indels 0; Gaps 0;									
QY	1	ATGAAGACCA	CATGTTTCTCTTATCAGTCTTATCTTAATTC	AAAGGGACAAAAATCTCCCC	60				
DB	1923	ATGAGACCA	CTGCCTGTTTCATCAGCTGATCTCTGATCCAGG	GCATCAGACCTCTGCC	1982				
QY	61	ATTTTAGAT	AGTAGCTAGTAATAATCAACCCAA	AAATGTGATGTCGCGAACT	120				
DB	1983	ATCTGAGAT	CGCAGCAACACACGCCCCAGAA	CGCTGTGACGGCACC	2042				
QY	121	CTCCAGA	GACAGACGCTTCATCTGATGGGATTCACA	CTGAGTGGGCAAAAGTTGCT	180				
DB	2043	CTGAGA	GAACCCAGGACGTGCACCTGATGGGCTTCA	CCCTGAGCGGCCAAGGTGCC	2102				
QY	181	GATTTC	CCCTTTGGAGGCATCCAA	CGCATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT	240				
DB	2103	GACAG	CCCTCTGGNGCCAGCAGAGTGGG	CCCTTCAGAGCCGGCTGCCCCCAGAC	2162				
QY	241	GTTGAGT	ACACAGAGGGGAGGAAGCAAA	CATGCTCAATATAAGTGTAA	CGGATCCC 300				
DB	2163	GTGGAGT	ACACCGGCGGAGGAGCCAGAC	CTGCTTACAACATCAGCTGACCGACCC	2222				
QY	301	TCGTGGA	AAATCCTTGTCTGTAGATCTCTTAC	CCNACATCCGTGACTATCCGAATGCAA	360				
DB	2223	AGCGG	CAAGAGCCTGTCTGTGACCCCTCC	CAACACATCAGGGACTACCTTAAGTGCAG	2282				
QY	361	ACTATC	CATCATATTCAAAGTCAAA	CCCTCATGSCACAGGGATCGCCCTTCATTTATGG	420				
DB	2283	ACCAT	CCACACATCCAGGSCCAGAC	CCCTCAAGCCAGGCGATCGCCCTGACCTGTGG	2342				
QY	421	GGAGCA	TTTTTTCTGTATATCGCATTTGCT	TCCACAAATGTACCGAGCAAAAGTCTTC	480				
DB	2343	GGCGC	CTTCTCTGTACGACAGATCG	CCAGCACCACTGTACAGGGCAGGGTGTTC	2402				
QY	481	ACTG	AGGGGAACATAGCAGCTATGATTGT	CAATAAGACAGTGCACAAAATGATTTCTCG	540				
DB	2403	ACCG	GGGCAACATCGCGGCATGATCG	TTTAAACAAGACCGGTGCACAAGATGATCTTCAGC	2462				
QY	541	CGGCA	AGGACAAAGGTACCGTCA	TATGAATCTGACTTCTACTATAATAATTGGACAGT	600				
DB	2463	AGGC	AGGCGCAGGGCTAAGGCA	CATGAACTGTACACAGCAACACAGTACTTGCAGCAGC	2522				
QY	601	AGTA	ACGGAAACGCAACGAATGAC	ACTGGATGTTTCGGCGCTCTTTCAAGAATACAATTC	660				
DB	2523	AACA	CGGACCCAGACCAACGAC	CACCGGCTGCTTCGGCGCCCTGCGAGGAGTACAACAGC	2582				
QY	661	ACAA	GAACCAAACATGTGCTCCGT	CCAAAATACCTCCACACTGGCCCAAGCCCGTCCG	720				

Db	2583	ACCAGNACAGACCTGCGCCCCCAGCGAAGATCCCGACGCCCTCGCCACCGCGCAGGCC	264
Qy	721	GAGATCAAACTCAACAGCACCCCAACTGATGCCAACCAAACTCAATACCGGACCCCAAGC	780
Db	2643	GAGATCAAGCCACCGACACCCCCACCGACGCCACCACTGAAACCAACCGACCCCAAC	2702
Qy	781	AGTGATGATGAGAGACCTCGCAACATCCGGCTCAGGGTCCGGAGNACGAGAAACCCACACA	840
Db	2703	AACGACGACGAGGACCTGTATCAACAGCGGCAGCGCGCAGCGAGCAGGACCCCTACACC	2762
Qy	841	ACTTCTGATCGGCTCAACAGCAAGGCGTTTCATCAACAATGCCACCCACTCCCTCACCAC	900
Db	2763	ACCAGCGAGCGCTGACCAAGCAGGCGCTGAGCAGCACCATGCTCTCTACCCCTAGCCCT	2822
Qy	901	CAACCAAGCAGCGCACGACGAAAGGAGGAAACAACAACCAATTCCTCCCAAGATGCTGTGACT	960
Db	2823	CAGCCGAGCACCCCTCAGCAGGAGGGGCAACAACAACCGACCAAGCCAGGCGACCGTGACC	2882
Qy	961	GAATAGACAAAATAACACAATGCAACAACCGTCCATGCCCTCATATAACACTTACCACA	1020
Db	2883	GAGCCCAACGAGACCAACACACCGCCGACCCAGCATGCTCTCTACAACACCCAGCC	2942
Qy	1021	ATCTCTACTAACACACCTCTCAAAACAACACTTCAGCACTCTCTCTGCAACCATTAACAAAC	1080
Db	2943	ATCAGCACCAACAACACCCAGCAAGAACAATTCAGCACCCCTGAGCGTGCGCTGCAGAAC	3002
Qy	1081	ACACCAATGACAAACACACAGAGACCAATCACTGAAAAATGAGCAAAACAGTGCCTCCCTG	1140
Db	3003	ACCACCAACTACGACACCCAGAGCACCGCCACCGAGAACGAGCAGCAGCGCCCTTAGC	3062
Qy	1141	ATAACAAACCTGCTCCCAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGCAAAAAA	1200
Db	3063	AGACCAACCTTGCTCCACCGGCACTTGACACACCGCAAGAGACCAACAACACCCAG	3122
Qy	1201	GGCCCCGCCCAACAGCGCACCAACACAGCAAAATGAGCATTTTACCAGTCTCCCTCCCAACC	1260
Db	3123	GGCCCCACCAACACCGGCCCTTAACATGACCAACGGCCACTGTACCAAGCCCGAGCCCAACC	3182
Qy	1261	CCGAGCTGACTGCACACATCTTGATATTTTTCAGAAAGCG	1304
Db	3183	CCCAACCCCAACCCAGCACCTGGTGACTTTCAGGAAGAAGAG	3226

RESULT 13

US-10-066-506A-11

; Sequence 11, Application US/10066506A

; Publication No. US20030108560A1

; GENERAL INFORMATION:

; APPLICANT: Grogan, Case C.

; APPLICANT: Hevey, Michael C.

; APPLICANT: Schmaljohn, Alan, L.

; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein

; FILE OF INVENTION: 003/243/SAP

; CURRENT APPLICATION NUMBER: US/10/066,506A

; PRIORITY DATE: 2002-01-31

; PRIOR FILING DATE: 2001-01-31

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Apple Macintosh Microsoft Word 6.0

; SEQ ID NO 11

; LENGTH: 2039

; TYPE: DNA

; ORGANISM: Ebola virus Zaire strain

; FEATURE:

; OTHER INFORMATION: chimeric molecule between Ebola virus Glycoprotein 1 and US-10-066-506A-11

Query Match 29.3%; Score 540; DB 15; Length 2039;
Best Local Similarity 99.1%; Pred. No. 3.5e-157;
Matches 543; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qv 1294 AGAAGAAAGCGATCGGCAATGTGTCATGTCTCAACCCAAATGCAACCCCTAATTTACATTAC 1353

```
Db 1492 AGAAGTCGAGATCGGCAATTTGCAATGCTCAACCCAAATGCAACCTTAATTTACATTAC 1551
Qy 1354 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTCGATACCATATTTTCGGGCA 1413
Db 1552 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTCGATACCATATTTTCGGGCA 1611
Qy 1414 GCAGCCGAGGGAATTTACATAGAGGGCTAAATGCACAATCAAGATGGTTAATCTGTGG 1473
Db 1612 GCAGCCGAGGGAATTTACATAGAGGGCTAAATGCACAATCAAGATGGTTAATCTGTGG 1671
Qy 1474 TTGAGACAGCTGGCCAAACGAGAGGACTCAAGCTCTTCAACTGTTCTGAGAGCCCAACT 1533
Db 1672 TTGAGACAGCTGGCCAAACGAGAGGACTCAAGCTCTTCAACTGTTCTGAGAGCCCAACT 1731
Qy 1534 GAGCTACGACCTTTTCAATCTCAACCGTAAGGCAATTTGATTTCTGCTGAGGATGG 1593
Db 1732 GAGCTACGACCTTTTCAATCTCAACCGTAAGGCAATTTGATTTCTGCTGAGGATGG 1791
Qy 1594 GCGGCACATGCCACATTTCTGGACCGGACTGCTGATCGAACCAACATGATGGACCAAG 1653
Db 1792 GCGGCACATGCCACATTTCTGGACCGGACTGCTGATCGAACCAACATGATGGACCAAG 1851
Qy 1654 AACATAACAGACAAATTTGATCAGATTTATCATGATTTTGTGATAAAACCTTCGGAC 1713
Db 1852 AACATAACAGACAAATTTGATCAGATTTATCATGATTTTGTGATAAAACCTTCGGAC 1911
Qy 1714 CAGGGGCAATGACAAATTTGGTGGACAGATGGAGCAATGATACCGGACGATTTGA 1773
Db 1912 CAGGGGCAATGACAAATTTGGTGGACAGATGGAGCAATGATACCGGACGATTTGA 1971
Qy 1774 GTTACAGGCGTTATATTCAGTTATCGCTTTATTTCTGATATGCAAAATTTCTTTTAG 1833
Db 1972 GTTACAGGCGTTATATTCAGTTATCGCTTTATTTCTGATATGCAAAATTTCTTTTAG 2031
Qy 1834 TTGAATTC 1841
Db 2032 TTGAATTC 2039
```

```
RESULT 14
US-10-860-878-4
; Sequence 4, Application US/10860878
; Publication No. US20050130129A1
; GENERAL INFORMATION:
; APPLICANT: Nabel, Gary J
; APPLICANT: Delgado, Rafael
; APPLICANT: Yang, Zhi-yong
; TITLE OF INVENTION: Targeting Gene Transfer Vectors To Certain Cell Types
; FILE REFERENCE: 2115a-001474POA
; CURRENT APPLICATION NUMBER: US/10/860,878
; PRIOR FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US/09/600,766
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: USSN 60/072033
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 7272
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Plasmid
; OTHER INFORMATION: containing DNA for soluble GP of Ebola Virus,
; OTHER INFORMATION: Zaire strain
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (886)..(1129)
; OTHER INFORMATION: CMV IE 5' UT
; FEATURE:
; NAME/KEY: intron
```

```
; LOCATION: (1130)..(1840)
; OTHER INFORMATION: CMV IE INT
; FEATURE:
; NAME/KEY: enhancer
; LOCATION: (248)..(885)
; OTHER INFORMATION: CMV enhancer
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1870)..(4288)
; OTHER INFORMATION: SGP (Z)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (4289)..(4841)
; OTHER INFORMATION: TSGH
; FEATURE:
; NAME/KEY: gene
; LOCATION: (6337)..(6959)
; OTHER INFORMATION: Kan r
; US-10-860-878-4

Query Match      28.7%; Score 527.6; DB 22; Length 7272;
Best Local Similarity 98.3%; Pred. No. 5.7e-153;
Matches 533; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1294 AGAAGAAAGCGATCGGCAATTTGTCAATGCTCAACCCAAATGCAACCTTAATTTACATTAC 1353
Db 3560 AGAATCGAAGAGAACAAATTTGTCAATGCTCAACCCAAATGCAACCTTAATTTACATTAC 3619
Qy 1354 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTCGATACCATATTTTCGGGCA 1413
Db 3620 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTCGATACCATATTTTCGGGCA 3679
Qy 1414 GCAGCCGAGGGAATTTACATAGAGGGCTTAATGCAACATCAAGATGGTTAATCTGTGG 1473
Db 3680 GCAGCCGAGGGAATTTACATAGAGGGCTTAATGCAACATCAAGATGGTTAATCTGTGG 3739
Qy 1474 TTGAGACAGCTGGCCAAACGAGAGGACTCAAGCTCTTCAACTGTTCTGAGAGCCCAACT 1533
Db 3740 TTGAGACAGCTGGCCAAACGAGAGGACTCAAGCTCTTCAACTGTTCTGAGAGCCCAACT 3799
Qy 1534 GAGCTACGACCTTTTCAATCTCAACCGTAAGGCAATTTGATTTCTGCTGAGGATGG 1593
Db 3800 GAGCTACGACCTTTTCAATCTCAACCGTAAGGCAATTTGATTTCTGCTGAGGATGG 3859
Qy 1594 GCGGCACATGCCACATTTCTGGACCGGACTGCTGATCGAACCAACATGATGGACCAAG 1653
Db 3860 GCGGCACATGCCACATTTCTGGACCGGACTGCTGATCGAACCAACATGATGGACCAAG 3919
Qy 1654 AACATAACAGACAAATTTGATCAGATTTATCATGATTTTGTGATAAAACCTTCGGAC 1713
Db 3920 AACATAACAGACAAATTTGATCAGATTTATCATGATTTTGTGATAAAACCTTCGGAC 3979
Qy 1714 CAGGGGCAATGACAAATTTGGTGGACAGATGGAGCAATGATACCGGACGATTTGA 1773
Db 3980 CAGGGGCAATGACAAATTTGGTGGACAGATGGAGCAATGATACCGGACGATTTGA 4039
Qy 1774 GTTACAGGCGTTATATTCAGTTATCGCTTTATTTCTGATATGCAAAATTTCTTTTAG 1833
Db 4040 GTTACAGGCGTTATATTCAGTTATCGCTTTATTTCTGATATGCAAAATTTCTTTTAG 4099
Qy 1834 TT 1835
Db 4100 TT 4101

RESULT 15
US-10-860-878-3
; Sequence 3, Application US/10860878
; Publication No. US20050130129A1
; GENERAL INFORMATION:
; APPLICANT: Nabel, Gary J
; APPLICANT: Delgado, Rafael
; APPLICANT: Yang, Zhi-yong
```

;; TITLE OF INVENTION: Targeting Gene Transfer Vectors To Certain Cell Types
;; TITLE OF INVENTION: By Pseudotyping With Viral Glycoprotein
;; FILE REFERENCE: 21158-001474FOA
;; CURRENT APPLICATION NUMBER: US/10/860,878
;; CURRENT FILING DATE: 2004-06-03
;; PRIOR APPLICATION NUMBER: US/09/600,766
;; PRIOR FILING DATE: 2001-05-14
;; PRIOR APPLICATION NUMBER: USSN 60/072033
;; PRIOR FILING DATE: 1998-01-21
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 3
;; LENGTH: 7285
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Plasmid
;; OTHER INFORMATION: containing DNA of GP of Ebola Virus, Zaire strain
;; FEATURE:
;; NAME/KEY: 5'UTR
;; LOCATION: (686)..(1129)
;; OTHER INFORMATION: CMV IE 5' UT
;; FEATURE:
;; NAME/KEY: Intron
;; LOCATION: (1130)..(1840)
;; OTHER INFORMATION: CMV IE INT
;; FEATURE:
;; NAME/KEY: enhancer
;; LOCATION: (1870)..(4301)
;; OTHER INFORMATION: GP(Z)
;; FEATURE:
;; NAME/KEY: gene
;; LOCATION: (4302)..(4854)
;; OTHER INFORMATION: TdGH
;; FEATURE:
;; NAME/KEY: gene
;; LOCATION: (6350)..(6972)
;; OTHER INFORMATION: Kan r
;; OTHER INFORMATION: Kan r
US-10-860-878-3

Query Match 28.7%; Score 527.6; DB 22; Length 7285;
Best Local Similarity 98.3%; Pred. No. 5.7e-153;
Matches 533; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1294 AGAAGNACCGATCGGCAATTGTCATGCTCAACCCAAATGCAACCCCTAATTACATTAC 1353
DB 3563 AGAACTCGAAGAGAGCAATTGTCATGCTCAACCCAAATGCAACCCCTAATTACATTAC 3622
QY 1354 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTGGATACCATATTTCGGGCCA 1413
DB 3623 TGGACTACTCAGGATGAAGGTGCTGCAATCGGACTGGCTGGATACCATATTTCGGGCCA 3682
QY 1414 GCAGCGAGGGAATTTACATAGAGGGCTAATGCAATCAAGATGGTTAATCTGTGGG 1473
DB 3683 GCAGCGAGGGAATTTACATAGAGGGCTAATGCAATCAAGATGGTTAATCTGTGGG 3742
QY 1474 TTGAGACAGCTGGCCNACGAGACGACTCAAGCTCTTCAACTGTTCTGAGAGCCCAACT 1533
DB 3743 TTGAGACAGCTGGCCNACGAGACGACTCAAGCTCTTCAACTGTTCTGAGAGCCCAACT 3802
QY 1534 GAGCTACGCACCTTTTCAATCTCAACCGTAAGGCAATTTGTTCTGCTGCGCGATGG 1593
DB 3803 GAGCTACGCACCTTTTCAATCTCAACCGTAAGGCAATTTGTTCTGCTGCGCGATGG 3862
QY 1594 GCGCGCACATGCCACATTTCTGGGACCGGACTGCTGTATCGAACCCATGATTGGACCAAG 1653
DB 3863 GCGCGCACATGCCACATTTCTGGGACCGGACTGCTGTATCGAACCCATGATTGGACCAAG 3922
QY 1654 AACATAACAGACAAAATTGATCAGATTATTATGATGTTTGTGATAAAACCCCTCCGGAC 1713

DB 3923 AACATAACAGACAAAATTGATCAGATTATTATCATGATTTTGTGATAAAACCCCTCCGGAC 3982
QY 1714 CAGGGGACCAATGACAAATTGGTGGACAGGATGGAGACAATGGATACCGGCAGGTATTGGA 1773
DB 3983 CAGGGGACCAATGACAAATTGGTGGACAGGATGGAGACAATGGATACCGGCAGGTATTGGA 4042
QY 1774 GTTACAGCGCTTATAAATTGCAAGTTATCGCTTTATTCTGTATATGCAAAATTTGCTTTTATAG 1833
DB 4043 GTTACAGCGCTTATAAATTGCAAGTTATCGCTTTATTCTGTATATGCAAAATTTGCTTTTATAG 4102
QY 1834 TT 1835
DB 4103 TT 4104

Search completed: August 1, 2005, 20:40:59
Job time : 1161.01 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:57:58 ; Search time 8793.18 Seconds
(without alignments)
11274.581 Million cell updates/sec

Title: US-10-066-506A-5
Perfect score: 2046
Sequence: 1 atgaagaccacattgttccct.....tcactaaatcatggatga 2046

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pi.*
9: gb_pr.*
10: gb_ro.*
11: gb_ats.*
12: gb_sv.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2046	100.0	2046	6	AX717717 Sequence
2	1841.2	90.0	11460	6	AR279677 Sequence
3	1841.2	90.0	19104	14	MYREPCYC
4	1841.2	90.0	19112	14	AY430366 Lake Vict
5	1841.2	90.0	19113	14	AY430365 Lake Vict
6	1813.6	88.6	2051	6	AX717721 Sequence
7	1695.6	82.9	2046	14	AF005735
8	1692.4	82.7	2948	14	MAVSPAA
9	1692.4	82.7	19112	14	MVIRPR
10	1653.6	80.8	19151	14	AY358025
11	1652.4	80.8	2046	14	AF005733
12	1518	74.2	2046	6	AX717725
13	1513.2	74.0	2046	14	AF005734
14	1319.6	64.5	2046	6	AX717719
15	1308	63.9	1841	6	AX717715 Sequence
16	546.6	26.7	2252	6	AX717713
17	148.6	7.3	2261	14	AY344234 Sudan ebo
18	148.6	7.3	2363	14	AY316199 Sudan ebo
19	148.6	7.3	18875	14	AY729654 Sudan ebo

20	145	7.1	2362	14	EVU23069
21	145	7.1	2362	14	EVU28134
22	138.8	6.8	2408	14	EVU28006
23	134.4	6.6	2408	14	EVU81161
24	132.8	6.5	18959	14	AY142960 Zaire Ebo
25	132.8	6.5	18960	14	AF499101
26	131.2	6.4	2039	6	AX717723 Sequence
27	131.2	6.4	2173	14	AY058898 Zaire Ebo
28	131.2	6.4	2298	6	AR404895
29	131.2	6.4	2298	6	AX092102 Sequence
30	131.2	6.4	2406	14	EVU31033
31	131.2	6.4	2408	14	EVU23187
32	131.2	6.4	2408	14	EVU28077
33	131.2	6.4	11742	14	EBORNA
34	131.2	6.4	18959	14	AF086833
35	131.2	6.4	18959	14	AF272001
36	131.2	6.4	18961	14	AY354458 Zaire ebo
37	129.6	6.3	2173	14	AY526098
38	129.6	6.3	2173	14	AY526100
39	129.6	6.3	2173	14	AY526101
40	129.6	6.3	2446	14	EVU77384
41	128	6.3	1041	14	AY526103
42	128	6.3	2173	14	AY526099
43	128	6.3	2173	14	AY526102
44	128	6.3	2173	14	AY526104
45	128	6.3	2173	14	AY526105

ALIGNMENTS

RESULT 1
AX717717
LOCUS
DEFINITION Sequence 5 from Patent WO02079239.
ACCESSION AX717717
VERSION AX717717.1 GI:29890727
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Grogan,C.C., Hevey,M.C. and Schmaljohn,A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 5 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source
1..2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chimeric molecule between Marburg virus strain Musoke Glycoprotein 1 and Marburg virus strain Raven Glycoprotein 2"

linear PAT 15-APR-2003

ORIGIN	Query Match	100.0%;	Score 2046;	DB 6;	Length 2046;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2046;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC	60		
Db	1	ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC	60		
Qy	61	ATTTTACAGATAGCTAGTAAATCAACCCCAAAATGGATTCCGTTATGTCCTCGGAAT	120		
Db	61	ATTTTACAGATAGCTAGTAAATCAACCCCAAAATGGATTCCGTTATGTCCTCGGAAT	120		
Qy	121	CTCCAGAACAGACAGAGCGTCCATCTCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	180		
Db	121	CTCCAGAACAGACAGAGCGTCCATCTCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	180		
Qy	181	GATTCCCTTTGGAGGAGATCCACCGATGGGCTTTCAGGACAGGTGTACTCTCCCAAGNAT	240		

```
Db      |||||
181  GATTCCTCCCTTTGGAGGCATCCAAGGCATGGGCTTTTCAGACAGAGTGTAACCTCCCAAGAA 240
Qy      |||||
241  GTTGTAGTACACAGAGGGGAGAGCAAAACATGCTCAAAATATAAGTGTACGGATCCC 300
Db      |||||
241  GTTGTAGTACACAGAGGGGAGAGCAAAACATGCTCAAAATATAAGTGTACGGATCCC 300
Qy      |||||
301  TCTGGAATAATCCTGTGTGTAGATCCTCTCAACATCCGTGACTATCCGAAATGCAAA 360
Db      |||||
301  TCTGGAATAATCCTGTGTGTAGATCCTCTCAACATCCGTGACTATCCGAAATGCAAA 360
Qy      |||||
361  ACTATCCATCATATTCAGGTCAAAACCCCTCATGACAGAGGGATCGCCCTTCATTTATGG 420
Db      |||||
361  ACTATCCATCATATTCAGGTCAAAACCCCTCATGACAGAGGGATCGCCCTTCATTTATGG 420
Qy      |||||
421  GGAGCATTTTTCTGTATGATCGCATGCTCCCAACAAATGTAACGAGGCAAGTCTTC 480
Db      |||||
421  GGAGCATTTTTCTGTATGATCGCATGCTCCCAACAAATGTAACGAGGCAAGTCTTC 480
Qy      |||||
481  ACTGAAGGGAAATAGCAGCTATGATTTGTCAATAAGACAGTGCACAAAATGATTTCTCG 540
Db      |||||
481  ACTGAAGGGAAATAGCAGCTATGATTTGTCAATAAGACAGTGCACAAAATGATTTCTCG 540
Qy      |||||
541  CGGCAAGCAAGGGTACCGTCAATATGAAATCTGACTTCTACTAAATAATTTGGACAAGT 600
Db      |||||
541  CGGCAAGCAAGGGTACCGTCAATATGAAATCTGACTTCTACTAAATAATTTGGACAAGT 600
Qy      |||||
601  AGTAACGGAAACCAACGAATGACACTGGATGTTTCGGGCGCTTCAAGAAATACAAATCT 660
Db      |||||
601  AGTAACGGAAACCAACGAATGACACTGGATGTTTCGGGCGCTTCAAGAAATACAAATCT 660
Qy      |||||
661  ACAAGAAACCAACATGTGCTCGTCCAAATACCTCCACCACTGCCACAGCCCGTCCG 720
Db      |||||
661  ACAAGAAACCAACATGTGCTCGTCCAAATACCTCCACCACTGCCACAGCCCGTCCG 720
Qy      |||||
721  GAGATCAAACTCAAGCAACCCAACTGATGCGCAACAACTCAATACACAGGACCCCAAGC 780
Db      |||||
721  GAGATCAAACTCAAGCAACCCAACTGATGCGCAACAACTCAATACACAGGACCCCAAGC 780
Qy      |||||
781  AGTGATGATGAGGACTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 840
Db      |||||
781  AGTGATGATGAGGACTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 840
Qy      |||||
841  ACTTCTGTGCTGCTCAACGAAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACCA 900
Db      |||||
841  ACTTCTGTGCTGCTCAACGAAGGGCTTTTCATCAACAAATGCCACCACTCCCTCACCA 900
Qy      |||||
901  CAACCAAGCAACGCAAGAGGAGAAACAAACAAACCAATTCGCAAGATGCTGTGACT 960
Db      |||||
901  CAACCAAGCAACGCAAGAGGAGAAACAAACAAACCAATTCGCAAGATGCTGTGACT 960
Qy      |||||
961  GAACTAGACAAAATAACACAACTGCAACCGTCCATGCGCCCTCATAACTACCAACA 1020
Db      |||||
961  GAACTAGACAAAATAACACAACTGCAACCGTCCATGCGCCCTCATAACTACCAACA 1020
Qy      |||||
1021  ATCTCTACTAACAACACCTCCAAACACAACTTTCAGCACTCTCTGCACTTACAAAC 1080
Db      |||||
1021  ATCTCTACTAACAACACCTCCAAACACAACTTTCAGCACTCTCTGCACTTACAAAC 1080
Qy      |||||
1081  ACCAACATGACACACAGAGGCAATCATCTGAAATGAGCAAAACGATGCGCCCTCG 1140
Db      |||||
1081  ACCAACATGACACACAGAGGCAATCATCTGAAATGAGCAAAACGATGCGCCCTCG 1140
Qy      |||||
1141  ATAAACAACCTGCTTCCACAGGGAATCCACACAGCAAGAGGACCCAGCAGCAAAA 1200
Db      |||||
1141  ATAAACAACCTGCTTCCACAGGGAATCCACACAGCAAGAGGACCCAGCAGCAAAA 1200
Qy      |||||
1201  GGGCCCGCCACCAACGGCACCACCAACGACCAATGAGCATTTCCACAGTCTCTCCCGCCACC 1260
Db      |||||
1201  GGGCCCGCCACCAACGGCACCACCAACGACCAATGAGCATTTCCACAGTCTCTCCCGCCACC 1260
Qy      |||||
1261  CCGAGCTCGACTGCAACAACTTTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAAA 1320
```

```
Db      |||||
1261  CCGAGCTCGACTGCAACAACTTTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAAA 1320
Qy      |||||
1321  GAAGGTGATATATTTCCCGTTTTAGATGGGTTAATAATACTGAAATTTGATTTGATCCA 1380
Db      |||||
1321  GAAGGTGATATATTTCCCGTTTTAGATGGGTTAATAATACTGAAATTTGATTTGATCCA 1380
Qy      |||||
1381  ATCCCAACACAGAAAACAACTTTTGTATGAAATCTCCAGCTTTAATACTTCAACTAATGAG 1440
Db      |||||
1381  ATCCCAACACAGAAAACAACTTTTGTATGAAATCTCCAGCTTTAATACTTCAACTAATGAG 1440
Qy      |||||
1441  GAAACAACACACTCCCGGAATATCAGTTTAACTTTCTCTATTTTCTCTGATAAAATGGA 1500
Db      |||||
1441  GAAACAACACACTCCCGGAATATCAGTTTAACTTTCTCTATTTTCTCTGATAAAATGGA 1500
Qy      |||||
1501  GATACTGCTACTCTCGGGGAAAACGAGAAATGATTTGTATGATGAGAGTTGAGGAGT 1560
Db      |||||
1501  GATACTGCTACTCTCGGGGAAAACGAGAAATGATTTGTATGATGAGAGTTGAGGAGT 1560
Qy      |||||
1561  GTGAGAGGAGACGATTTTGGCGGAGGCTTAGCTGAGTACCAATTTTGGCCCTCGGAATC 1620
Db      |||||
1561  GTGAGAGGAGACGATTTTGGCGGAGGCTTAGCTGAGTACCAATTTTGGCCCTCGGAATC 1620
Qy      |||||
1621  GAAGGACTCTATATCTGCGGTTTTAATCAAAATCAGAAACAATTTAGTTTGTAGTTGAGG 1680
Db      |||||
1621  GAAGGACTCTATATCTGCGGTTTTAATCAAAATCAGAAACAATTTAGTTTGTAGTTGAGG 1680
Qy      |||||
1681  CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGTCACAAACGAGGAA 1740
Db      |||||
1681  CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGTCACAAACGAGGAA 1740
Qy      |||||
1741  AGGACATTTTCTTAATCAATAGGATGCAATTTGATTTTGTCTTACGAGTGGGGCGGA 1800
Db      |||||
1741  AGGACATTTTCTTAATCAATAGGATGCAATTTGATTTTGTCTTACGAGTGGGGCGGA 1800
Qy      |||||
1801  ACATGCAAGGCTGATGAGACCTGATTTGTGATAGGAATAGAGATCTATCTAAAAATATC 1860
Db      |||||
1801  ACATGCAAGGCTGATGAGACCTGATTTGTGATAGGAATAGAGATCTATCTAAAAATATC 1860
Qy      |||||
1861  TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAAGGAGGAACTGGCTGGGGCTTA 1920
Db      |||||
1861  TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAAGGAGGAACTGGCTGGGGCTTA 1920
Qy      |||||
1921  GGTGGCAAAATGGTGGACATCTGACTGGGGTGTCTCACCAATTTGGGCAATCTCTGACTA 1980
Db      |||||
1921  GGTGGCAAAATGGTGGACATCTGACTGGGGTGTCTCACCAATTTGGGCAATCTCTGACTA 1980
Qy      |||||
1981  TTATCTATAGCTGTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db      |||||
1981  TTATCTATAGCTGTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Qy      |||||
2041  GGATGA 2046
Db      |||||
2041  GGATGA 2046

RESULT 2
AR279677
LOCUS      AR279677              11460 bp      DNA      linear      PAT 10-APR-2003
DEFINITION Sequence 1 from patent US 6517842.
ACCESSION  AR279677
VERSION    AR279677.1      GI:29714600
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 11460)
AUTHORS   Hevey,M.C., Negley,D.L., Pushko,P., Smith,J.F. and Schmaljohn,A.L.
TITLE     Marburg virus vaccines
JOURNAL   Patent: US 6517842-A 1 11-FEB-2003;
          Location/Qualifiers
FEATURES   source
            1..11460
            /organism="unknown"
            /mol_type="genomic DNA"
```

ORIGIN

Query Match 90.0%; Score 1841.2; DB 6; Length 11460;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTCTCTATAGTCTTATCTTAATTTCAAGGACAAAAATCTCCC 60
DB |||||
QY 5940 ATGAAGACACATGTTCTCTATAGTCTTATCTTAATTTCAAGGACAAAAATCTCCC 5999
DB |||||

QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGTATGCTCCGGAAT 120
DB |||||

QY 6000 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGTATGCTCCGGAAT 6059
DB |||||

QY 121 CTCGAGAGACAGAGAGCTCATCTGATGGATTCACATGAGTGGGCAAAAGTTGCT 180
DB |||||

QY 6060 CTCGAGAGACAGAGAGCTCATCTGATGGGATTCACATGAGTGGGCAAAAGTTGCT 6119
DB |||||

QY 181 GATTCCTCTTTGGAGGCATCAAGGATGGGCTTTTCAGACAGGTGTACCTCCCAAGAT 240
DB |||||

QY 6120 GATTCCTCTTTGGAGGCATCAAGGATGGGCTTTTCAGACAGGTGTACCTCCCAAGAT 6179
DB |||||

QY 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAAATATAGTGAACGGATCCC 300
DB |||||

QY 6180 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAAATATAGTGAACGGATCCC 6239
DB |||||

QY 301 TCTGAAAAATCTTCTGTGTAGATCCTCTACCAATCCGTGACTATCCGAAATGCAAA 360
DB |||||

QY 6240 TCTGAAAAATCTTCTGTGTAGATCCTCTACCAATCCGTGACTATCCGAAATGCAAA 6299
DB |||||

QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
DB |||||

QY 6300 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 6359
DB |||||

QY 421 GGAGCATTTTCTGTATGATCGCATGCTCCCAACAAATGATGACAGGCAAAAGTCTTC 480
DB |||||

QY 6360 GGAGCATTTTCTGTATGATCGCATGCTCCCAACAAATGATGACAGGCAAAAGTCTTC 6419
DB |||||

QY 481 ACTGAAGGAAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAATGATTTCTCG 540
DB |||||

QY 6420 ACTGAAGGAAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAATGATTTCTCG 6479
DB |||||

QY 541 CGGCAAGGACAGGGTACCGTATATGATCTGACTTCTAATAAATATTTGGCAAGT 600
DB |||||

QY 6480 CGGCAAGGACAGGGTACCGTATATGATCTGACTTCTAATAAATATTTGGCAAGT 6539
DB |||||

QY 601 AGTAAACGGAACCAACGAAATGACATGATGTTTCGGCGCTTTCAAGAAATACAAATCT 660
DB |||||

QY 6540 AGTAAACGGAACCAACGAAATGACATGATGTTTCGGCGCTTTCAAGAAATACAAATCT 6599
DB |||||

QY 661 ACAAGAAACCAACATGTGCTCCGTCCTCAAAATACCTCCACCACTGCCACAGCCGTCGG 720
DB |||||

QY 6600 ACAAGAAACCAACATGTGCTCCGTCCTCAAAATACCTCCACCACTGCCACAGCCGTCGG 6659
DB |||||

QY 721 GAGATCAAACTCAAGACCCCACTGATGTCACCAAACTCAATACCAAGCCCAAGC 780
DB |||||

QY 6660 GAGATCAAACTCAAGACCCCACTGATGCCACCAAACTCAATACCAAGCCCAAGC 6719
DB |||||

QY 781 AGTGTATGATGAGGACCTCGCAACATCCGCTCAGGGTCGGAGAACGAGAACCCCAACACA 840
DB |||||

QY 6720 AGTGTATGATGAGGACCTCGCAACATCCGCTCAGGGTCGGAGAACGAGAACCCCAACACA 6779
DB |||||

QY 841 ACTTCTGATGCGGTCAACGAAGAGGCTTTTCATCAACAAATGCCACCACTCCCTCACCA 900
DB |||||

QY 6780 ACTTCTGATGCGGTCAACGAAGAGGCTTTTCATCAACAAATGCCACCACTCCCTCACCA 6839
DB |||||

QY 901 CAAACGAACGCGCACAGCAAGAGGAAACAAACAAACCATTTCCCAAGATGCTGTGACT 960
DB |||||

QY 6840 CAAACGAACGCGCACAGCAAGAGGAAACAAACAAACCATTTCCCAAGATGCTGTGACT 6899
DB |||||

QY 961 GAACTAGACAAAATAACCAACTGACAAACCGTCCATGCGCCCTCATAACTACCACTACCACA 1020
DB |||||

QY 6900 GAACTAGACAAAATAACCAACTGACAAACCGTCCATGCGCCCTCATAACTACCACTACCACA 6959
DB |||||

QY 1021 ATCTTACTAAACAACACTCTCAAAACAACACTTTCAGCACTCTCTCTGCAACCATTAACAAC 1080
DB |||||

QY 6960 ATCTTACTAAACAACACTCTCAAAACAACACTTTCAGCACTCTCTCTGCAACCATTAACAAC 7019
DB |||||

QY 1081 ACCACCAATGACAAACACACAGAGCACAATCACTGAAAATGAGCAAAACCAAGTCCCTCG 1140
DB |||||

QY 7020 ACCACCAATGACAAACACACAGAGCACAATCACTGAAAATGAGCAAAACCAAGTCCCTCG 7079
DB |||||

QY 1141 ATAAACAACCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAACAGAGCAAAAAA 1200
DB |||||

QY 7080 ATAAACAACCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAACAGAGCAAAAAA 7139
DB |||||

QY 1201 GGCCCGCCCAACGGGACCAAAACAGCAAAATGAGCAATTTCAACAGTCTCTCCGCCAC 1260
DB |||||

QY 7140 GGCCCGCCCAACGGGACCAAAACAGCAAAATGAGCAATTTCAACAGTCTCTCCGCCAC 7199
DB |||||

QY 1261 CCCAGCTCGATGCAACAACATCTTGTATATTTTCAGAACAAAGCGATTCGATTTCTGGAAA 1320
DB |||||

QY 7200 CCCAGCTCGATGCAACAACATCTTGTATATTTTCAGAACAAAGCGATTCGATTTCTGGAGG 7259
DB |||||

QY 1321 GAAGTGTATATATTTCCCGTTTTTATAGTGGTTAAATACTGAAAATTTGATTTGATCCA 1380
DB |||||

QY 7260 GAAGTGTATATATTTCCCGTTTTTATAGTGGTTAAATACTGAAAATTTGATTTGATCCA 7319
DB |||||

QY 1381 ATCCAAAACACAGAAAACATCTTTGATGATCTCCAGCTTTAATACTTCAACTAATGAG 1440
DB |||||

QY 7320 GTTCCAAAATACAAAACAAATCTTTGATGATCTCTAGTCTCTGGTGGCTCGGCTGAGGAA 7379
DB |||||

QY 1441 GAAACAACACATCTCCCGGAATATCACTTTTAACTTTCTCTATTTTCTGATAAAAAATGGA 1500
DB |||||

QY 7380 GATCAACATGCTCTCCCAATATTTAGTTTAACTTTATCTTATTTTCTTAATAAATGAG 7439
DB |||||

QY 1501 GATCTGCTACTCTCTGGGAAAAACGAGAAATGATGTGATGACAGAGTTGAGGATTTGGAGT 1560
DB |||||

QY 7440 AACCTGCTACTCTCTGGAGAAAATGAGAAATGATGTGATGACAGAGTTAAGAAATTTGGAGC 7499
DB |||||

QY 1561 GTGAGGAGGACGATTTGCGCGCAGGCTTACGTGATGATACCAATTTTGGCCCTCGAATC 1620
DB |||||

QY 7500 GTTTCAGGAGGATGACCTGCGCGCAGGCTCAGTTGGATACCGTTTTTTTGGCCCTCGAAT 7559
DB |||||

QY 1621 GAAGGACTCTATCTACCTGCGGTTTTAATAAATAACAGAAATTTAGTTTGTAGTTTGGG 1680
DB |||||

QY 7560 GAAGGACTCTATCTACCTGCTGTTTTAATAAATAACAGAAATTTAGTTTGTAGTTTGGG 7619
DB |||||

QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTTTGGAGCTCTTTGTTAAGGCTCACACAGGAA 1740
DB |||||

QY 7620 CGCTTAGCTAATCAAACTGCTAAATCTTTGGAGCTCTTTGTTAAGGCTCACACAGGAA 7679
DB |||||

QY 1741 AGGACATTTTCTTAACTAATAGGACATGCAATGATCTTTTGTCTTACGAGGTGGGGCGGA 1800
DB |||||

QY 7680 AGAACATTTCTCTTAACTAATAGGACATGCAATGATCTTTTGTCTTACGAGGTGGGGCGGA 7739
DB |||||

QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTTGATAGGAATAGAGATCTATCTAAAAATATC 1860
DB |||||

QY 7740 ACATGCAAGGTGCTAGGACCTGATTTGTTGATAGGAATAGAGATCTTGTCCAAAAATAT 7799
DB |||||

QY 1861 TCAGAACAAATCGACAAAATCAGAAAGGATGAAACAAAAAGGAGGAACTGGCTGGGGTCTA 1920
DB |||||

QY 7800 TCAGAACAAATCGACAAAATCAGAAAGGATGAAACAAAAAGGAGGACCTGGTGGGGTCTG 7859
DB |||||

QY 1921 GGTGCAAAATGGTGACATCTGACTGGGGTGTTCACAAATTTGGGCACTCTGTACTA 1980
DB |||||

QY 7860 GGTGTAATGGTGACATCTGACTGGGGTGTTCCTACTAATCTGGGCACTTTGTACTA 7919
DB |||||

QY 1981 TTATCTATAGTGTCTGATGCTCTGCTGATCTGCTGATCTCTTCAATAAATACAT 2040
DB |||||

QY 7920 TTATCTATAGTGTCTGATGCTCTGCTGATCTGCTGATCTCTTACTAAATATATC 7979
DB |||||

QY 2041 GGATGA 2046
DB |||||

QY 7980 GGATGA 7985
DB |||||

Gene	APENPLNGFNRRQVLLAYANPTLSAV"
mRNA	5822..8666 /gene="gp" 5822..8666 /gene="gp" /standard_name="gene four" /citation={3} /evidence=experimental 5940..7985 /gene="gp" /standard_name="gp" /codon_start=1 /citation={3} /evidence=experimental /product="glycoprotein" /protein_id="CAA78117.1" /db_xref="GI:60627" /db_xref="GOA:P35253" /db_xref="UniProt/Swiss-Prot:P35253" /translation="MKTTCFLISLILQGTKNLPILFIETASNNQPNQVDSVCSGTLQRT EDVHLMGFTLSQKVDAPLEAKSRWAPRTGVPPKKNVEYTEGEAKTCYNI SVTPDSG KSLLDPPNTRIDYPKCTIHHI OQONPHAGIALHLWGAFFLYDR IASTITMRGVKF TEGNTAMLVNKHMI PSQOGQYRHMNLSTNKYWTSSNGTQNDTGCFGALQBY NESTKQTCAPSKIPPLTPAPEIKLSTPTDITAKLNTDPSDEDLATSSGSGSR EPHTTSQVKSLSMPPPTSPQSPQOQGNTHNSQDAVTELDKNNITTAQSNP PHNTTITSTNTSKINFATLAPLQNTNDNTQSTI TENEQTSAPSTLTPPTGNPT AKSTSGKGPATNTNEHTSPPTSPSTAOHLVFRKRSLILWREGDMFPFLDG LINAPIDPVPNTKTI PDESSSGASAEQHASPNISLTSVFPNINENTAYSGEN ENDCAELIRHSVQEDDLAAGLSWIPFPFGIEGLIYTA VLIKNQNLVCLRLANQT AKSLELLURVTEERTFSLNRHAI DFLITRWGGTCKVLGPDCCIGIEDLSKNISEQI DOIKDEQKRGWGSLGKWNWTSQWGLTNLIGLLLSIAVLIALSCICRIFTKYIG" 8762..10009 /gene="vp30" 8762..10009 /gene="vp30" /standard_name="gene five" /citation={3} /evidence=experimental 8864..9697 /gene="vp30" /standard_name="vp30" /citation={3} /codon_start=1 /evidence=experimental /product="vp30" /protein_id="CAA78118.1" /db_xref="GI:541782" /db_xref="GOA:P35258" /db_xref="UniProt/Swiss-Prot:P35258" /translation="MQQPRGRSTRNHQVPTIYHETQLPSKPHYTNVHPRRSMST RSASRSPNTHI PRAPPPSTNLSKPPPPKDMCRNMKIGLPCADPTCNDRDLDNLT NRELLLLMARQMLPNTDTKFRMPQCGSPSLSKGLSKDQKQDTVDLTLENLHLSY LHRSEIGNMWRHLRAALSITCAGIRKTRNSLINTLHMHENLPDQNGVTKQYTT GTHLDKGGQFEAALMQGWDKRSISLIFQAALYVNNIPCESSISIVQASYESPYSSKS R" 9990..11278 /gene="vp24"
CDS	Query Match 90.0%; Score 1841.2; DB 14; Length 19104; Best Local Similarity 93.7%; Pred. No. 0; Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0; 1 ATCAAGACACATGTTTCCTTATCATGCTCTATCTTAATTCAGGGACAAAAATCTCCCC 60 5940 ATGAAGACACATGTTTCCTTATCATGCTCTATCTTAATTCAGGGACAAAAATCTCCCC 5999 61 ATTTTAGAGTAGTACTAGTAATCAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 120 6000 ATTTTAGAGTAGTACTAGTAATCAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 6059 121 CTCAGAAGACAGAAGACGTCATCTGATGGGATTCACATCGAGTGGGCAAAAAGTTGCT 180 6060 CTCAGAAGACAGAAGACGTCATCTGATGGGATTCACATCGAGTGGGCAAAAAGTTGCT 6119

Query Match	90.0%;	Score 1841.2;	DB 14;	Length 19104;
Best Local Similarity	93.7%;	Pred. No. 0;		
Matches 1918;	Conservative	0;	Mismatches 128;	Indels 0; Gaps 0;
1	ATCAAGACCACATGTTTCCCTTATCAGTCTTATCTTAATTCAAGGGACACAAAAATCTCCCC	60		
5940	ATGAGACCACATGTTTCCCTTATCAGTCTTATCTTAATTCAAGGGACACAAAAATCTCCCC	5999		
61	ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGTCCTCGGAAC	120		
6000	ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGTCCTCGGAAC	6059		
121	CTCCAGACAGACAGAACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAATGTGCT	180		
6060	CTCCAGACAGACAGAACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAATGTGCT	6119		

QY	1261	CCAGCTCGACTGCAACAACATCTTGTAATATTTTCAGAGAAGCGATCGATTTCTTGGAAA	1320
Db	7200	CCAGCTCGACTGCAACAACATCTTGTAATATTTTCAGAGAAGCGAAGTATCTCTGGAGG	7259
QY	1321	GAAGTGATATATCCCGTTTTAGATGGTTAAATAATACATAAATGATTTGATCCCA	1380
Db	7260	GAAGGCGACATGTTCCCTTTCTCGATGGTTAAATAATGATTTGATTTGATCCCA	7319
QY	1381	ATCCCAACACAGAAACAATCTTTGATGAATCTCCAGCTTTAAATACCTTCAACTAATGAG	1440
Db	7320	GTTCCTCAATACAAACAATCTTTGATGAATCTCTAGTCTGGTCCCTCGCTGAGGNA	7379
QY	1441	GAACAACACATCCCGCGAATACGATTTAACTTTCTTATTTCTGATATAAATGGA	1500
Db	7380	GATCAACATGCTCCCGCAATATAGTTTAACTTTATCTTATTTCTTAAATAATGAG	7439
QY	1501	GATAGCTCTACTGCGGGAACGAGATGATGTCGACAGTTGAGGATTTGGAGT	1560
Db	7440	AACACTGCTCTACTGCGGGAACGAGATGATGTCGACAGTTGAGGATTTGGAGC	7499
QY	1561	GTGCGAGGAGCGATTTGGCGGAGGCTTAGCTGGATACCAATTTTGGCCCTCGAATC	1620
Db	7500	GTTCAGGAGATGACTGCGCGAGGCTCAGTTGGATACCGTTTGGCCCTCGAAT	7559
QY	1621	GAAGGACTCTATGCTCGCGGTTTAAATCAAAAATCAGAACAAATTTAGTTGTAGTTGAGG	1680
Db	7560	GAAGGACTTTACACTGCTGCTTTTAAATTAATAATCAAAACAATTTGCTGTCAGGTTGAGG	7619
QY	1681	CGCTTAGCTTAATCAAACTGCTAATCTTGGAGCTTCTTTAAGGTCACACCGGAGNA	1740
Db	7620	CGTCTAGCCAAATCAAACTGCCAAATCTTTGGAACTCTTATTGAGAGTCACAACTGAGGA	7679
QY	1741	AGGACATTTCTTAAATCAATAGGATGCAATTTGACTTTTCTTACGAGGTGGGGCGGA	1800
Db	7680	AGAACATTTCTTAAATCAATAGATGATGCTATTGACTTTCTACTCACAAGATGGGGAGA	7739
QY	1801	ACATGCAAGGTCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTTAAAAATATC	1860
Db	7740	ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGATAGAGACTTTGTCCAAAAATAT	7799
QY	1861	TCAGAACAAATCGACAAATCAGAAAGATGAACAAAGGAGGAACTGGCTGGGGTCTA	1920
Db	7800	TCAGAGCAAAATGACAAATTAATAAGGACGACAAAGAGGGGACTGGTTGGGGTCTG	7859
QY	1921	CGTGGCAAAATGCTGGACATCTGACTGGGGTGTCTCACCAGTTTGGGCACTCTGCTACTA	1980
Db	7860	GCTGGTAAATGCTGGACATCCGACTGGGGTGTCTTACTTACTTGGGCACTTGTGCTACTA	7919
QY	1981	TTATCTATAGCTGTTCTGATGCTCTGCTCTATCTGTCGATCTTCACTAAATACATT	2040
Db	7920	TTATCTATAGCTGTTCTGATGCTCTGCTCTATCTGTCGATCTTCTTACTAAATATATC	7979
QY	2041	GGATGA 2046	
Db	7980	GGATAA 7985	
RESULT 4			
AY430366			
LOCUS	19112 bp	RNA	linear VRL 06-FEB-2004
DEFINITION	Lake Victoria marburgvirus strain pp4 guinea pig nonlethal variant, complete genome.		
ACCESSION	AY430366		
VERSION	AY430366.1	GI:40388379	
KEYWORDS	Lake Victoria marburgvirus		
SOURCE	Lake Victoria marburgvirus		
ORGANISM	Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.		
REFERENCE	1 (bases 1 to 19112)		
AUTHORS	Hevey, M., Negley, D., Geisbert, J., Jahrling, P., and Schmaljohn, A.		
TITLE	Antigenicity and vaccine potential of Marburg virus glycoprotein expressed by baculovirus recombinants		

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

Virology 239 (1), 206-216 (1997)
98087840
9426460
2 (bases 1 to 19112)
Chain, P.S.G., Malfatti, S.A., Hajjaj, A., Vergez, L.M., Do, L.H.,
Smith, K.L., and McCreedy, P.M.
Direct Submission
Submitted (09-OCT-2003) Viral Sequencing Group, Lawrence Livermore
National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA
3 (bases 1 to 19112)
Ichou, M.A., Paragas, J., Jahrling, P.B., Ibrahim, M.S., Lofes, L.,
Hevey, M., and Schmaljohn, A.
Direct Submission
Submitted (09-OCT-2003) Viral Genomics Group, USAMRIID, Fort
Detrick, MD 21702, USA
Location/Qualifiers

source

1. 19112

/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="pp4 guinea pig nonlethal variant"
/isolate="Musoke Kenya, 1980"
/db_xref="taxon:11269"
104..2191
/gene="NP"
3
/gene="NP"
/note="nucleoprotein"
/codon_start=1
/product="NP protein"
/protein_id="AAR85453.1"
/db_xref="GI:40388380"
/translation="MDLSLLELGTPTAPHRNKKVILFDTHQVSIQNQIIDAIS
GIDNGLLEGLLTLCEHYNSDKFNTSPIAKYLADAGYEDVIEKADATRLDV
IPNPHYSPLTLALKTLESTSORGRIGFLSLPLKLVGDRAGIEKALROVTV
HQEGIVTYPNHLLTGHMKVIFGLISSFLIKVLJHGVNLVTHGDAYSILISNV
QTFRSGLLIVLEFILQKTDGVTLLPLVRISKVKNVASFQALSNLARHGEYA
PFARVNLGKNNLEHGLYPQLSAIALGVAHTGSTLAGVNVGEYQQLREAAHDAEV
KLQRHHEQEIQAIAEDDEERKILEQPHLQKTEITHSQTLAVLSQKREKLARAAEIE
NNIVRDEQFKQSONRVSQFLNDPTPEVTQVQARPMNRPTALPPVDDKIBHES TEDS
SSSSFDLNDPFDLLEDEDTLDDSVMI PCTTSREFOGIPPEPQSGDLNNSQCKOE
DESTNPIKFLRYQELPPVQEDDESEYTTDSQSIDQPGSDNEGVDLPPPLAQE
KRQPIQHPANQDPFGSDGVNDGIDLEIRSPSSPAQEDTKMRGAYELSPDTN
DEDNQNPQPVVTKGRTFLYPLNDLQTNPPESLITLALVEEYQNPVSAKELQADWPD
MSFDRRHVAMNL"

gene

CDS

104..2191

/gene="NP"

104..2191

/note="NP"

/codon_start=1

/product="NP protein"

/protein_id="AAR85453.1"

/db_xref="GI:40388380"

/translation="MDLSLLELGTPTAPHRNKKVILFDTHQVSIQNQIIDAIS
GIDNGLLEGLLTLCEHYNSDKFNTSPIAKYLADAGYEDVIEKADATRLDV
IPNPHYSPLTLALKTLESTSORGRIGFLSLPLKLVGDRAGIEKALROVTV
HQEGIVTYPNHLLTGHMKVIFGLISSFLIKVLJHGVNLVTHGDAYSILISNV
QTFRSGLLIVLEFILQKTDGVTLLPLVRISKVKNVASFQALSNLARHGEYA
PFARVNLGKNNLEHGLYPQLSAIALGVAHTGSTLAGVNVGEYQQLREAAHDAEV
KLQRHHEQEIQAIAEDDEERKILEQPHLQKTEITHSQTLAVLSQKREKLARAAEIE
NNIVRDEQFKQSONRVSQFLNDPTPEVTQVQARPMNRPTALPPVDDKIBHES TEDS
SSSSFDLNDPFDLLEDEDTLDDSVMI PCTTSREFOGIPPEPQSGDLNNSQCKOE
DESTNPIKFLRYQELPPVQEDDESEYTTDSQSIDQPGSDNEGVDLPPPLAQE
KRQPIQHPANQDPFGSDGVNDGIDLEIRSPSSPAQEDTKMRGAYELSPDTN
DEDNQNPQPVVTKGRTFLYPLNDLQTNPPESLITLALVEEYQNPVSAKELQADWPD
MSFDRRHVAMNL"

gene

CDS

2945..3934

/gene="VP35"

2945..3934

/gene="VP35"

/note="minor component of RNA-dependent RNA-polymerase"

/codon_start=1

/product="VP35 protein"

/protein_id="AAR85454.1"

/db_xref="GI:40388381"

/translation="MWDSSYMQQVSEGLMTGKVPIDQVFGANPLKLYKRRPKGTVG
LQSPCLMSKATSDDDIIMDOLI VKRTLADLLI PINROI SDIQTSLSEVTRVHEIER
QLHEITPVLMKGRTELEAISKGMSLEMLAKYDLVITSTRTTAPAAAFDAYLNEHGVPPP
OPAIKFDLQVAQACSKGTWVKNATTDAAADKMSKVLSEETFSKPNLSAKDLALLLF
THLGNNTPPHILAQVLSKIAYKSGSGAFDAPHOILSEGENAQAALTRLSRTDFAF
LGVVPPVTRVKNFQTVPRPCQKSLRAVPPNPITDKGVCVYSSEQGETRALKI"

gene

CDS

4568..5479

/gene="VP40"

4568..5479

/gene="VP40"

/note="matrix protein"

/codon_start=1

/product="VP40 protein"

/protein_id="AAR85455.1"

/db_xref="GI:40388382"

/translation="MASSSNNTYMOYLNPPPYADHGANQLIPADOLSNQOQGITPNVY
GDNLDDQFKGNVCHATLEAI IDISA VNETVKGVPAWPLGIMSNSEYPLAHTVAA
LLTGSYITQTTHNGQKFPVRNLTGTGIPAPLMLREGNQAFIQNWVIPNFSNQF
TYNTNLVLSVQKLPDDARPSKDKLIGNTHHPAVSIHPNLPPIVLPVTKQAYKQHK
PNNGPLLAISGIIHLQRVEKPEKTSIFRISLSPADMFPSVKEGMMKRGENSEPVVYFQ


```

/db_xref="GI:40388388"
/translation="MDLSLLELGTGKTPAIVRVKKNVILFTNHOVSIQNIIDAINS
GIDGLDEGLLTLVCHYNSDKRSTPIAKYLRDAGTEFVQKADATFLDV
IPNEPHYSPLIALKLTUESTESQRIIGLFLSCFLPQLVWGDRIASIEKARQVTV
HQEQIVTPYNNHLLTTHMKVIFGLIRSSFLKFLVLIHQGNVLTGHDAYDIIISNV
QTRFSGLLIVKTVLEFLQKDSGVTLHPLVRTSKVNEVASFOALSILARHGEVA
PFARVLNSGILNLEHGLYPOLSAIALGVATAHGSTLAGVNVGROYOQLREAAHDAEV
KLQRHHEQEIQAEDDEERKILEQFHLQKTEITHSOTLAVLSQKREKLARLAAEIE
NNIVEDQFGSQNRVSFLNDPTPEVTVQARPMNKPITALPPVDDKILHESHTDS
SSSSFDLNDPFPALLNEDETLDDSVMI PGTTSRFQGIPEPPRQSDLNNSQKQE
DESTNPKQLFRLYQELFPVQEDDESEYTTDSQESIDQPSDNEQGVDLPPPLIYAE
KRQDPIQHPAANPODPFGSIGDVNGDILEPIRSPSSPAQEDTRMREAYELSDFDFT
DEDNQONWFORVTVKKGRTFLYPNDLLQTNPESLITALVEEYQNPVSAKELQADWPD
MSFDERHVAANL"
2945..3934
/gene="VP35"
CDS
/db_xref="GI:40388389"
/translation="WDSSSYMQQVSEGLMTGKVPIDQVFGANPLKLYKREKPKGTVG
LQCSPLMSKATSDDIINDQLIVKRTILADLLIPINQISDIQSTLSEVTVRVEIER
QHEITTPVLMKGRITLAEISKGMSEMLAKYIDHLVISTGRTTAPAAAFDAYLNEHGVPVP
QPAIFKDIGVQAQCSGTGKNATDAADKMSVLESEETFSKPNLSAKDLALLIF
THLGNNTPFHILAQVLSKIAYKSGKGAFLDAFHQLISEGENAQAALTRLSTFDAP
LGVVPVIRVKNFQTVPRPCQKSLRAVPPNPTIDKGWCVVYSSEQGETRALKI"
4568..5479
/gene="VP40"
CDS
/db_xref="GI:40388390"
/translation="MASSSYNTYQYLNPPYPADHGANQLIPADQLSNOQGITPNYV
GLLNDQFGKNVCHAFTELEAIIDISAYNERTVGVPAWLPLGMSFEYPLAHTVAA
DLTGSYTTIQTHNGQKVRVNRILGTGIPAPHLRLMREGNOAFTQNNVI PRNFSTNOF
TNNLTNLSVQKLPDDAWRPSKNKLI GNTMHPAVSIHPNLPPIVLPTVKQAYROHK
NPNGPLAISGILHQLRVEKVPKTSLSRISLPADMFVSVEGMKRGKSPVYVYQ
APENPLNGFNRRQVVLAYANPTLSAV"
5941..7986
/gene="GP"
CDS
/db_xref="GI:40388391"
/translation="WKTCTFLISLILQGTKNLPILEIASNNQPNQVDSVCSGLQKT
EDVHLMPGLSGQVADSPLASKRAFRTGVPNNVETGEERAKTCINISVTDPSG
KSLLDPTNTRIDVPKCTIHHIQGNPHQAQIALHLWGAFPLDYDIASTMTYRGKVP
TEGNIAMVANKTVHMKIFRQQGYRHMLTSTNKYVTSNGSDTQDTCFGALQEY
NSTNOTCAPSKI PPLPTARPEIKLSTPTDACKLNTDPSDDGDLATSGSGGER
EPHTTSDAVTQGLSMTWPTTSPQSPPTQCGNNTNHSQDAVTELDKNNTTAQPSM
PHNTTISTNNTSKNFSTLSAQLQNTNDNTQSTITENEQTSAPSIITLPTPTNPT
AKSTSKGPNATPATNTEHFTSPPTTSPQAHLVFRKRSILMLREGMDPFLDG
LINAPIDPDPVNTKTIIDESSSGASAEQDQASPNISLTLSPFNINENTAYSGEN
ENDCDAEIRIWSVEDDILAAGLSWIPFGPGIEGLYTAGLIGNNNVILNRRRLANOT
AKSLELLAVTTEERTFSLNRHAI DFLLRWGGTKVGLPDCGIGEDLSKNTSEQI
DQIKDEQEGTGWGLGKWKVTDGWLVTNLGILLLLLSIAVLIALSCICRIFTYIG"
8869..9714
/gene="VP30"
CDS
/db_xref="GI:40388392"
/translation="MAELSTRYNLPANVTENSINLDNSTARWIKBSVSGWTVKMGN
FVPHIPNTGMTLHLHKSFWFPEWQOTRNLFSHLFPNPKSTIIEPLALRILGLVAL
KDQELQOQLIPGRFSIVHLSWELLLEVTSAIHLSPNLLGIYLTDMF KILMAGVNF
FNKMTLHVVDHDKGSSIRIKLTGQOIITRVMGFLVEVRRIDIRPCCGCTVLSBS
VWGLVAEAVLREHSQMEKQGPLNLQYMSKIAL"
11481..18476
/gene="L"
CDS
/db_xref="GI:40388393"
/translation="MAELSTRYNLPANVTENSINLDNSTARWIKBSVSGWTVKMGN
FVPHIPNTGMTLHLHKSFWFPEWQOTRNLFSHLFPNPKSTIIEPLALRILGLVAL
KDQELQOQLIPGRFSIVHLSWELLLEVTSAIHLSPNLLGIYLTDMF KILMAGVNF
FNKMTLHVVDHDKGSSIRIKLTGQOIITRVMGFLVEVRRIDIRPCCGCTVLSBS
VWGLVAEAVLREHSQMEKQGPLNLQYMSKIAL"
11481..18476
/gene="L"
CDS
/db_xref="GI:40388394"
/translation="MQHPTQVPDARLSPILLQCDLLARSILGYSYSHNPKLRNCR
IPHIYRLNSTALKTFLQNCSTLTVPHSITWDHILTSIOYDANHVDDFKYLLPSEL
RKYANWNEFLKAYLNKILGLDHFVSASARSQCEDFSPKENPYTWQMLLLVHLQSLAR
RIGQRGLSRNSNMKFIGTLEFGIADVFIPKPKTII RNNAVSLQAGKGLRIWYRD
QNTYPLCDDDEFIVSADSECFINIKDVFIERVNTWBEICARAWLEDGADYPPDLVL
GELYNGDQIITAMYLEDFKLIKHLBPCVSIQTHGIFTPRKYTFWFSQMIKSYDEL
HDNLKLIQISDNKAECAQNFITKIVQAKLTPQOYCELSKQKHGHPVLYNDVALDKV
KHAQSTKILPKVMFETFCVEKFI VAKNHVSQGSWKYTHDLHLTPYLROHTVSNV
PFSQAEIYQHLMEWYFVEHEPLSTKIIISLISFIKDRATAVQECDSVDFRVLGY
NPVPRQSVRVPQGLQADFSLNQILEFAKLEYLAPSTYFNALFSKEKELNLGRIF
GKPYRVNRQVPTLAEALLADGLAKAPFNNMMVVTERRQKALLHOASWHHNSASIGEN
AIVRGASVDFLEKINLAFYEFYTHFIDYCNQGVKNFLPDMWHFIPLCYMHVSDF
YSPHCYTEDNRNPPDCANAYHYHGLGIECQYKQWTCISCAOITIVELKTKLKS
CWMGDNOCITLTSIFPIDAPNDVQENAEALNARVAVELAITTCVSGISGIFLAKPETFVH
SGFIYFKQKQYLVGQVLPQSLTKMARGCPISDSIFDDLOQSLSIGTSFERTSETRH
IPFSRWIASQHYMLNLLNQNLHGFPLGNFIDISCFKQPLTSFSEKIALITPQVLGG
LSFLNPEKLFYRNISDPLTSGFQKNALFEKEBELFYILISKPGGLADSDFVMP
LGLNPGSKEIITFLRQTVRENTITQSNRIINSLFHIGSDLEQVCEWILLSNPVM
SRPAADIFSRTPSGKRLQVLGLEGTRTLAASRTISLTTEGTMKMLRELTRNRKWS
FSYDALDDLSLSLEXETCTVDVANFLRAYSWSDVLKKGELIGATLPCLEQPEVKW
INLSDEQFNSSDSKSTINLLPYDCKELRLEGSDNTELNYSVSCALDRKVQKHPS
VNRLAWTIGNRAPYIGSRTEKIGYPLRVNCPASAALKEAIBWVSRLMLWVTQADRE
KLITPLNRSVNDYQTNLPLPHTYSGNIVHRYNDQYQGHSPFMANRMSNTSTRAIIS
TNTLKGVAGGQAAIDSNII FQNTINLGVAVLDIALSLAKSSASNTVFLMLNKCCT
RHYPSYLYFDKPLDVNLKYMDELVDNDPLCSGIGRIGRVSRSTLSLSLVSDI
GSDPPTIAWTLGETIVGSI FSDSSQSTDPISSGCTKTFVTHLYPVVPSIYVAG
ANLIVLSURKISIKNSDLTILISSTIRNLSHRSLRILQSTFRHELVTPLRAHHI
PLTSLMGGAGKESSDAVRLFTASYNQFINNPSCLMKKQSSLPWLYVPSBEGQ
KLPKTLKQSLDLSLSPDKIQKRLADTCCPIGFSFWYPSKTRHLYASLAWRD
KANKVYTPSHSLINCSFPBSHSTSVSSNOQVTSKYIYVPMNI TEINARTLINY
GSTALOQDMKPLSEONLVCNCPSEGI RPKDNQKITKHDRCERESSPQOMFED
NMQTPAHHSPPQIILKSLDAHEDFASKIILNSEINNLTEYLTNTKLTPTPR
TELDTPSLOSRYSSTRSLLSREQASIVYDCSNIPISLDYVSSNQVQ
MLNITYKRDYLFACDSNOFCFTGVSSMHYKLYDLLPPGKLKAI CLAEEGSGARL
LLKWKETDLPFTLATDSQOEAELSGRIVPRMLYNI DRLSLAKKALREIENNLTIQ
IITVDLWLDSDIYQVLPEDSDILTMDAETTKDETREOLYKTI VNIWTRTSPNPKIS
IIKVLADYEGSTLFLMKNALQYGVQKPYSSNAKNSWYLCGCKGRLOIOLDFIS
DOVGIFLICKAMSRQRAIIPYWLKHIEKNYPASLHEFFLTPGFSLESSEFCHRTIIF
SEGALPHKQSVYRQKQHLHLSMLDYENNSPLDLNRHFI CSIRKATIKTYINDILK
LNLVKAVERGKNWSQVBEILPNMHSCI VHVDEHSCGCEKRLLLKLDITRNTKIAEQ
KLLNRVGYILFPFGLFKGSLRA"

```

ORIGIN

Query Match 90.0%; Score 1841.2; DB 14; Length 19113;
 Best Local Similarity 93.7%; Pred. No. 0;

Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;			
QY	1	ATGAAGACACATGTTTCTCTATCAGTCTTATCTTAAATTCAGGGACAAAAAATCTCCCC	60
Db	5941	ATGAAGACACATGTTTCTCTATCAGTCTTATCTTAAATTCAGGGACAAAAAATCTCCCC	6000
QY	61	ATTTTATGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	120
Db	6001	ATTTTATGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	6060
QY	121	CTCCAGAGACAGAGAGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTGTCT	180
Db	6061	CTCCAGAGACAGAGAGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTGTCT	6120
QY	181	GATTCCTCTTGGAGGATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT	240
Db	6121	GATTCCTCTTGGAGGATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT	6180
QY	241	GTGTAGTACACAGAGGGGAGGAAGCCAAAAATCATCTCAATATATAGTGTAAACGATCCC	300
Db	6181	GTGTAGTACACAGAGGGGAGGAAGCCAAAAATCATCTCAATATATAGTGTAAACGATCCC	6240
QY	301	TCGTGAAATCTTCTGCTGTAGATCCTCTCAACATCCCGTGAATCCGGAATGCAAA	360
Db	6241	TCGTGAAATCTTCTGCTGTAGATCCTCTCAACATCCCGTGAATCCTTAAATGCAAA	6300
QY	361	ACTATCCATCATATTCAGGTCAAAACCTCATGACACAGGGGATCGCCCTTCATTTATGG	420
Db	6301	ACTATCCATCATATTCAGGTCAAAACCTCATGACACAGGGGATCGCCCTTCATTTATGG	6360
QY	421	GGAGCATTTTTCTGTATGATCGCATTCCTCCACAAATGATACCGAGGCAAAAGTCTTC	480
Db	6361	GGAGCATTTTTCTGTATGATCGCATTCCTCCACAAATGATACCGAGGCAAAAGTCTTC	6420
QY	481	ACTGAAGGAAACATAGCAGCTATGATGTCAATATAGACAGTGCAAAATGATTTCTCG	540
Db	6421	ACTGAAGGAAACATAGCAGCTATGATGTCAATATAGACAGTGCAAAATGATTTCTCG	6480
QY	541	CGGCAAGGACAGGGTACCGTCAATGATCTGACTTCTACTTAATAATATTTGGACAAGT	600
Db	6481	CGGCAAGGACAGGGTACCGTCAATGATCTGACTTCTACTTAATAATATTTGGACAAGT	6540
QY	601	AGTAAACGGAACGCAACGAAATGACATGGATGTTTTCGGGCTCTTCAAGAAATACAAATCT	660
Db	6541	AGTAAACGGAACGCAACGAAATGACATGGATGTTTTCGGGCTCTTCAAGAAATACAAATCT	6600
QY	661	ACAAAGAACCAAAACATGTGCTCGGTCCAAATACTCTCAGCACTGCGCACAGCCCGTGG	720
Db	6601	ACAAAGAACCAAAACATGTGCTCGGTCCAAATACTCTCAGCACTGCGCACAGCCCGTGG	6660
QY	721	GAGTCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACACGGACCCCAAGC	780
Db	6661	GAGTCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACACGGACCCCAAGC	6720
QY	781	AGTGTATGATGAGGACCTCGCAACATCCGCTCAGGGTCCGGAGAACGAGAACCCCAACA	840
Db	6721	AGTGTATGATGAGGACCTCGCAACATCCGCTCAGGGTCCGGAGAACGAGAACCCCAACA	6780
QY	841	ACTTCTGATGCGGTACCAAGCAAGGGCTTTCAATCAAACTGCGACCCCTCCCTCAGCA	900
Db	6781	ACTTCTGATGCGGTACCAAGCAAGGGCTTTCAATCAAACTGCGACCCCTCCCTCAGCA	6840
QY	901	CAACCAAGCACCCACAGCAGGAGGAAACACAAACCAATTCCTCCAGATGCTGTGACT	960
Db	6841	CAACCAAGCACCCACAGCAGGAGGAAACACAAACCAATTCCTCCAGATGCTGTGACT	6900
QY	961	GAACTAGACAAAAATAACAACATGCAACACCGTCCATGCCCCCTCATAACTACTACCACA	1020
Db	6901	GAACTAGACAAAAATAACAACATGCAACACCGTCCATGCCCCCTCATAACTACTACCACA	6960
QY	1021	ATCTCTACTAAACACACCTCCAAAACAACTTTAGCAGCTCTCTCTGCGCACTTACAAAC	1080
Db	6961	ATCTCTACTAAACACACCTCCAAAACAACTTTAGCAGCTCTCTCTGCGCACTTACAAAC	7020

RESULT 6
AX717721
LOCUS

AX717721

2051 bp

DNA

linear

PAT 15-APR-2003


```

DEFINITION   Sequence 9 from Patent WO02079239.
ACCESSION    AX717721
VERSION      AX717721.1  GI:29890729
KEYWORDS
SOURCE
ORGANISM     Marburg virus (strain Musoke)
              Marburg virus (strain Musoke)
              Viruses; ssRNA negative-strand viruses; Mononegavirales;
              Filoviridae; Marburg-like viruses.
1
REFERENCE
AUTHORS      Grogan,C.C., Hevey,M.C. and Schmaljohn,A.L.
TITLE        Chimeric filovirus glycoprotein
JOURNAL      Patent: WO 02079239-A 9 10-OCT-2002;
              U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
             Location/Qualifiers
             1..2051
             /organism="Marburg virus (strain Musoke)"
             /mol_type="unassigned DNA"
             /db_xref="taxon:33727"
             /note="chimeric molecule between Marburg virus
             Glycoprotein 1 and Marburg virus Glycoprotein 2"

ORIGIN
Query Match      88.6%; Score 1813.6; DB 6; Length 2051;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 1910; Conservative 0; Mismatches 124; Indels 12; Gaps 1

Qy 1 ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAAATTCAGGGACAAAAAATCTCCCC 60
Db 10 ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAAATTCAGGGACAAAAAATCTCCCC 69
Qy 61 ATTTTAGAGTAGCTAGTAAATAACAACCCAAAATGTGGATTCGGTATGCTCCGGAATCT 120
Db 70 ATTTTAGAGTAGCTAGTAAATAACAACCCAAAATGTGGATTCGGTATGCTCCGGAATCT 129
Qy 121 CTCGAGAGACAGAAAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTCGT 180
Db 130 CTCGAGAGACAGAAAGCGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTCGT 189
Qy 181 GATTCCCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAAAT 240
Db 190 GATTCCCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACTCCCAAGAAAT 249
Qy 241 GTTGAGTACACAGAGGGGGAGAGCCAAAACATGCTACAATAAAGTGAACGGATCCC 300
Db 250 GTTGAGTACACAGAGGGGGAGAGCCAAAACATGCTACAATAAAGTGAACGGATCCC 309
Qy 301 TCTGGAAAATCCTTGCTGTAGATCCTCTCTACCAACATCCCGTACTATCCGAAATGCAAA 360
Db 310 TCTGGAAAATCCTTGCTGTAGATCCTCTCTACCAACATCCCGTACTATCCGAAATGCAAA 369
Qy 361 ACTATCCATCATATTCAAGGTCAAAACCCTCATGCACAGGGGATCGCCTTCATTATGG 420
Db 370 ACTATCCATCATATTCAAGGTCAAAACCCTCATGCACAGGGGATCGCCTTCATTATGG 429
Qy 421 GGAGCATTTTTTCTGTATGATCGCATTCGCTCCACAAATATACCGAGGCAAAAGTCTTC 480
Db 430 GGAGCATTTTTTCTGTATGATCGCATTT-----ATGTACGGAGCAAAAGTCTTC 477
Qy 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATAAGACAGTGCAAAAATGATTTTCTCG 540
Db 478 ACTGAAGGGAACATAGCAGCTATGATTTGTCATAAGACAGTGCAAAAATGATTTTCTCG 537
Qy 541 CGGCAAGGACAGAGGTGACCGTCTATGATCTGACTTCTACTATAATAATATTGGACAAGT 600
Db 538 CGGCAAGGACAGAGGTGACCGTCTATGATCTGACTTCTACTATAATAATATTGGACAAGT 597
Qy 601 AGTAAACGGAACGCAAAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATCAAAATCT 660
Db 598 AGTAAACGGAACGCAAAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATCAAAATCT 657
Qy 661 ACAAGAACCAAAACATGTGCTCCGTCCAAAATACCTTCCACCACTGCCACAGCCCGTCGG 720
Db 658 ACAAGAACCAAAACATGTGCTCCGTCCAAAATACCTTCCACCACTGCCACAGCCCGTCGG 717

```

```
QY 1801 ACATGCAAGGCTCTAGGACCTGATTGTTGTCATAGGAATAGAAAGATCTATCTTAAATAATATC 1860
Db 1798 ACATGCAAGGCTCTAGGACCTGATTGTTGTCATAGGAATAGAAAGATCTATCTTAAATAATATC 1857
QY 1861 TCAGAACAAATCGACAAATCAGAAAGATGAAACAAAGGAGGAACTGGCTGGGGTCTTA 1920
Db 1858 TCAGAACAAATGACAAATTAAGGAGGACGAAACAAAGGAGGAGCTGGTGGGGTCTG 1917
QY 1921 GGTGGCAATGGTGACATCTGACTGGGGTGTCTCACCAGATTTGGGCATCTGCTACTA 1980
Db 1918 GGTGGTAAATGGTGACATCCGACTGGGGTGTCTTACTACTTGGGCATTTTGTACTA 1977
QY 1981 TTAATCTATAGCTGTTCTGAATGCTCTGCTCTGATATCTGTCGATCTTCACTAAATACAT 2040
Db 1978 TTAATCTATAGCTGTTCTGAATGCTCTGCTCTGATATCTGTCGATCTTACTAAATATATC 2037
QY 2041 GGATGA 2046
Db 2038 GGATAA 2043

RESULT 7
AF005735          2046 bp      RNA      linear      VRL 23-JAN-1998
Marburg virus strain M/Germany/Marburg/1967/Ratayczak glycoprotein
precursor (GP) gene, complete cds.
ACCESSION      AF005735
VERSION
KEYWORDS
SOURCE
ORGANISM
Lake Victoria marburgvirus
Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
1 (bases 1 to 2046)
Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and
Feldmann,H.
Variation in the glycoprotein and VP35 genes of Marburg virus
strains
Virology 240 (1), 138-146 (1998)
JOURNAL
MEDLINE
PUBMED
98110148
948698
2 (bases 1 to 2046)
Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
Direct Submission
Submitted (29-MAY-1997) Special Pathogens Branch, Centers for
Disease Control and Prevention, 1600 Clifton Road, Building 15,
Room S8611, Mail Stop G14, Atlanta, GA 30333, USA
Location/Qualifiers
1. .2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/Germany/Marburg/1967/Ratayczak"
/db_xref="taxon:11269"
1. .2046
/gene="GP"
/gene="GP"
/notes="virion surface glycoprotein"
/codon_start=1
/product="glycoprotein precursor"
/protein_id="AAC40460.1"
/db_xref="GI:2459880"
/translation="MKTKLFLSLILIOGKTLPILEIASNNPONVDSVCSGTLQKT
EDVHLMGPTLSQKVADSPLEASKEWAFRTGVPKKNVETGEERAKCYNLSVTDPSG
KSLILDPPNIDYPKCKTHIHIOQNPHAQSIALHLNGAFPLDYRIASTWYGRVFP
TEGNTAAMLVNKTVMHMFPSROGQRHNLTSFNKWTSSNNGTQNDTGCFALQDEY
NSTKNQTCAPSKIPSLPTARPEIKPTSTPTDATTNTDQNNDEDLITSGSGSEQ
EPYTTSDAVTKQGLSMTWPPSPQPSPTQOEGNNDHSDQVTVEPNKNTNTAQSPMP
PHNTTASNTNKNFTLSVSLQNTNYDQSTATENEQTSAPSKTLLPTGNLT
AKSTNTKGPPTTANMTNGHLLTSPSPNPTTQHLVFRKRSILWREGDMFPLDG
LINAIFDPPVENTKTIIDESSSSGASAEEDOHASPNISLTLSPFNINENTAYSGEN
ENDCDAELRWISVQSDLLAAGLUSWIPFGPGLEGYTAGLIGNKNQLVCRLLRANQT
AKSLELLRLRVTTTEERTFSLINRHAIIDFLLTRWGGTCKVLGPDCCIGIEDLSRNTSEQI
```

```
DOIKDQKEGTGMLGKWTSDWGLVTLNLGILLILLSIAVLIALSCICRIPTKYG"
1. .81
/gene="Gp"
82. .2043
/gene="Gp"
/product="glycoprotein"
1930. .2019
/gene="Gp"
/notes="encodes transmembrane anchor peptide"

ORIGIN

Query Match      82.9%; Score 1695.6; DB 14; Length 2046;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 1827; Conservative 0; Mismatches 219; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTCTTATCAGTCTTATCTTAATCAAGGACAAAAAATCTCCCC 60
Db 1 ATGAAGACCAATGTTCTTATCAGTCTTATCTTAATCAAGGATAAAAAATCTCCCT 60
QY 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGAACT 120
Db 61 ATTTTAGAGATAGCTAGTAATCAATCAACCCCAAAATGTGGATTCGGTATGCTCCGAACT 120
QY 121 CTCAGAAAGACAGAAAGCTTCATCTGATGGATTCACTGAGTGGGCAAAAATTTGCT 180
Db 121 CTCAGAAAGACAGAAAGATGTCCTATCTGATGGGATTCACACTGAGTGGGCAAAAATTTGCT 180
QY 181 GATTCCCTTTGGAGGCATCCAAGCGATCCAGGCTTTTCAGACAGAGGTGTACCTCCCAAGAA 240
Db 181 GATTCCCTTTGGAGGCATCCAAGCGATCCAGGCTTTTCAGACAGAGGTGTACCTCCCAAGAA 240
QY 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAATGTTAAACGGATCCC 300
Db 241 GTTGAGTATACAGAGGGGAGGAGCCAAACATGCTACAATATAATGTTAAACGGATCCC 300
QY 301 TCTGAAAAATCTTGTCTGTTAGATCCTCTTACCAACATCCGTGACTATCCGNAATGCAAA 360
Db 301 TCTGAAAAATCTTGTCTGTTAGATCCTCTTACCAACATCCGTGACTATCCGNAATGCAAA 360
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
QY 421 GGAGCATTTTTTCTGTATGATCGCATTCCTCCCAACAATGTATCCGAGGCAAAAGTCTTC 480
Db 421 GGAGCATTTTTTCTGTATGATCGCATTCCTCCCAACAATGTATCCGAGGCAAAAGTCTTC 480
QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCTCG 540
Db 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCTCG 540
QY 541 CGGCAAGGACAGAGGTACCGTCATATGAATCTGACTTCTACTATAATAATATTGGACAAGT 600
Db 541 AGGCAAGGACAGAGGTACCGTCATATGAATCTGACTTCTACTATAATAATATTGGACAAGT 600
QY 601 AGTAAACGGAACGCAAAACGAATGACATCGGATGTTTTCGGCGCTTCTTCAAGAATACAAATCT 660
Db 601 AACAAATGGAACACAAAACGAATGACATCGGATGTTTTCGGCGCTTCTTCAAGAATACAAATCT 660
QY 661 ACAAGAACCAAAACATGTGCTCCGTCCAAAATACCTCCACCACTCCCCACAGCCCGTCCG 720
Db 661 ACGAAGAAATCAAAACATGTGCTCCGTCCAAAATACCTCCACCACTCCCCACAGCCCGTCCG 720
QY 721 GAGATCAAACTCACAAGCACCCCAACTGATGCGCCCAAACTCAATATCCACGAGCCCAAGC 780
Db 721 GAGATCAAAACCAACAGCACCCCAACTGATGCGCCCACTCACTACACACAGCCCAAGC 780
QY 781 AGTGTATGATGAGGACCTCCGAACATPCGGCTCAGGGTCCGGGAAACGAGAAACCCACACA 840
Db 781 AATGTATGATGAGGACCTCTCAATCAATCCGGTTCAGGGTCCGGGAAACGAGAAACCCATACA 840
QY 841 ACTTCTGATGGGTGACCAAGCAAGGCTTTTCATCAACATGCCACCCCACTCCCTCACCA 900
```

Db 841 ACTTCAGATGCGGTCACTAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA 900
Qy 901 CAACCAAGCAGCCCAACAGCAGGAGGAAACAACAACCAATTCCTCCCAAGATGCTGTGACT 960
Db 901 CAACCAAGCAGCCCAACAGCAGGAGGAAACAACAACCAATTCCTCCCAAGATGCTGTGACT 960
Qy 961 GAACCTAGACAAAATAACACAACTGCACAAACGCTGCATGCCCTCATACCACTACCAACA 1020
Db 961 GAACCAACAAACCAACAACAATCTGCAACACGCTGCATGCCCTCCCAACAACCACTGCA 1020
Qy 1021 ATCTCTACTAACAACACCTCCAAACCAACAATTTAGCACTCTCTCTGACCACTTACAAAC 1080
Db 1021 ATCTCTACTAACAACACCTCCAAAGAACAACTTCAGCACCTCTCTCTGATCACTACAAAC 1080
Qy 1081 ACCACAATGACACACAGAGGACAAATCACTGAAATAGAGCAAAACAGTGCCCTCCG 1140
Db 1081 ACCACAATGACACACAGAGGACAACTGAAATAGAGCAAAACAGTGCCCTCCG 1140
Qy 1141 ATAACAACCTGCTCCCAACGGGAATCCACACAGCAAGAGACACAGCAGCAAGAAA 1200
Db 1141 AAAACAACCTGCTCCCAACAGGAATCTTACACAGCAAGAGACACAAACACAGGAAA 1200
Qy 1201 GGGCCCGCACAAACGGCACCAACCAAGATGAGCAATTCACCAAGTCCCTCCCGCAC 1260
Db 1201 GGGCCCGCACAAACGGCACCAATATGACAAATGGCAATTAACCAAGTCCCTCCCGCAC 1260
Qy 1261 CCCAGCTGAGTGACAAACATCTCTATATTTTCAAGAAAGCGATCGAATTTCTGGAAA 1320
Db 1261 CCCAACCCGACACACAAACATCTCTATATTTTCAAGAAAGCAAGTATCTCTGGAGG 1320
Qy 1321 GAAGTGATATATTTCCGTTTTAGATGGTTTAAATATGATGATGATGATGATGATGATGAT 1380
Db 1321 GAAGTGATATATTTCCGTTTTAGATGGTTTAAATATGATGATGATGATGATGATGATGAT 1380
Qy 1381 ATCCCAACACACAGAAACAACTTTGATGAATCTCCAGCTTTAATATCTCAACTAATGAG 1440
Db 1381 GTTCCAAATACAAAGACGATCTTGAATGATCTTCTAGTCTGGTCTTCGGCTGAGAA 1440
Qy 1441 GAACCAACACATCCCGCAATATCAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1500
Db 1441 GATCAACATGCTCCCGCAATATCAATTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1500
Qy 1501 GATACGCTACTCTGGGGAACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 AACATGCTACTCTGGGGAACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Qy 1561 GTTCAGGAGGAGATTTGGCGGAGGCTTAGCTGATACCAATTTTGGCCCTCGAATC 1620
Db 1561 GTTCAGGAGGAGATTTGGCGGAGGCTTAGCTGATACCAATTTTGGCCCTCGAATC 1620
Qy 1621 GAAGCACTTATATCTGCGGTTTAAATCAAAAATCAAGAAATTTAGTTTGTAGTTGAGG 1680
Db 1621 GAAGCACTTATATCTGCGGTTTAAATCAAAAATCAAGAAATTTAGTTTGTAGTTGAGG 1680
Qy 1681 CGCTTAGCTAATCAAACTGCTAAATCTTTGGAGCTCTTGTAAAGGTCACAAACGAGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTAAATCTTTGGAGCTCTTGTAAAGGTCACAAACGAGAA 1740
Qy 1741 AGGACATTTCTTAAATCAATAGGATCAATTTGATCTTTTGTCTTACAGGTTGGGCGGA 1800
Db 1741 AGGACATTTCTTAAATCAATAGGATCAATTTGATCTTTTGTCTTACAGGTTGGGCGGA 1800
Qy 1801 ACATCAAGGTTCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAAATATC 1860
Db 1801 ACATCAAGGTTCTTGGACCTGATTTGTCATAGGAATAGAGATCTTCCAGGAATAT 1860
Qy 1861 TCAGAAACAAATCGACAAATCGAAGGATGAAACAAAGGAGGAACTGGCTGGGGCTCTA 1920
Db 1861 TCAGAAACAAATCGAAGGATGAAACAAAGGAGGAACTGGCTGGGGCTCTA 1920
Qy 1921 GGTGGCAATGTTGACATCTGACCTGGGGTGTCTCACCAATTTGGGATCTCTGCTACTA 1980
Db 1921 GGTGGTAAATGTTGACATCTGACCTGGGGTGTCTTACTAACTTGGGCAATTTTGCTACTA 1980

Qy 1981 TTATCTATAGCTGTCTGATTGCTCTGCTGCTGATCTGCTGATCTTCACTAAATACATT 2040
Db 1981 TTATCCATAGCTGTCTGATTGCTCTGCTGCTGATCTTGTGATCTTTTACCAAAATATATC 2040
Qy 2041 GGATGA 2046
Db 2041 GGGTAA 2046
RESULT 8
MAVSPAA
LOCUS Marburg Virus genomic RNA of GP gene. linear VRL 22-JAN-1996
DEFINITION X68493
ACCESSION X68493
VERSION GI:296960
KEYWORDS gp gene; structural protein.
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE
1 (bases 1 to 2948)
Bukreyev, A., Volchkov, V.E., Blinov, V.M. and Netesov, S.V.
The GP-protein of Marburg virus contains the region similar to the
'immunosuppressive domain' of oncogenic retrovirus P15E proteins
FEBS Lett. 323 (1-2), 183-187 (1993)
JOURNAL
MEDLINE 93265932
PUBMED 8495737
REFERENCE
2 (bases 1 to 2948)
Bukreyev, A.A., Volchkov, V.E., Blinov, V.M., Dryga, S.A. and
Netesov, S.V.
The complete nucleotide sequence of the Popp (1967) strain of
Marburg virus: a comparison with the Musoke (1980) strain
Arch. Virol. 140 (9), 1589-1600 (1995)
JOURNAL
MEDLINE 96028047
PUBMED 7487490
REFERENCE
3 (bases 1 to 2948)
Bukreyev, A.A.
Direct Submission
TITLE
Submitted (28-SEP-1992) A.A. Bukreyev, All-Union Inst. of Molecular
Biology, Location/Qualifiers
FEATURES
source
1..2948
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="Popp"
/db_xref="taxon:11269"
5..2851
/gene="GP"
/gene_start=1
/product="structural protein"
/protein_id="CAA48507.1"
/db_xref="GI:296961"
/db_xref="GOA:P35254"
/db_xref="UniProt/Swiss-Prot:P35254"
/translation="MKVTCLEPISLILQIKTLPILEIASNNQPNQVDSVCSGTLOKT
EDVHLMGFTLSGQKVDSPLEAKRWAFRTGVPNPNVEYTGSEAKTCINISVDPFG
KSLLDLPPTNIRDYPKCTIHIIQGNPHAGIHLWGAFLYDRIAISTMTYGRVF
TEGNIAMIVNKTVMKIFSRQGYRHNLTSTNKTWTSNNGTQNDTNDGFGALQEV
NSTKNQCAPSKI PSLPTARPEIKPTSTPTDATTANTDNDDEDLITSGSGSEQ
EPYTTSDAVTKQGLSMTPTPSPOPSTPOEGNNTDHSQCTVTEPNKNTNTAOPSM
PHNTALSTNNTSKNFSTLSVSLQNTNTDQSTATENEQTSAPSKTLLPTGNLTT
AKSTNNTKGPPTTAPNMTNGLTSPSPNPTQHLVYFRKRSILMREGDMFPFLDG.
LINAPIDFPVPNTKTIYFDESSSSGASABEDHASPNSILSLTSLFPNNINFTYSGEN
ENDCDAELRWVQEDDLAAGLSWIPFPGIEGLYTAGLIGNKNLVCRLRLANQ
AKSLELLRVTTTEERTFSLINRHAIDFLTWGCTCKVLGPDCCIGIEDLRNITSEQI
DQIKKRBQKSGTGWLGKWTSDWGLVTNLGILLLLSLVIALSCICRIPTKYIG"

ORIGIN

Query Match		82.7%;	Score 1692.4;	DB 14;	Length 2948;
Best Local Similarity		89.2%;	Pred. No. 0;		
Matches 1825;		Conservative	0;	Mismatches	221;
		Indels		0;	Gaps
				0;	
QY	1	ATGAAGACCAATGTTCTCTTATCAGTCTTATCTTAATTCAGGACACAAAATCTCC	60	DB	1142
DB	122	ATGAAGACCAATGTTCTCTTATCAGTCTTATCTTAATTCAGGACACAAAATCTCC	181	QY	1081
QY	61	ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGCGATTGCGGTATGCTCCGGA	120	DB	1202
DB	182	ATTTTAGAGATAGCTAGTAACAAATCAACCCCAAAATGCGATTGCGGTATGCTCCG	241	QY	1141
QY	121	CTCCAGAGACAGAGAGCGTCCATCTGATGGGATTCACATGAGTGGGCAAAAGTTG	180	DB	1262
DB	242	CTCCAGAGACAGAGAGTGTCCATCTGATGGGATTCACATGAGTGGGCAAAAGTTG	301	QY	1200
QY	181	GATTCCTCTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTTACCTCCCAAG	240	DB	1321
DB	302	GATTCCTCTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTTACCTCCCAAG	361	QY	1260
QY	241	GTGTAGTACACAGAGGGGAGGAGCCAAAACATGCTACAATATAAGTGTAAACGG	300	DB	1381
DB	362	GTGTAGTATACAGAGGGGAGGAGCCAAAACATGCTACAATATAAGTGTAAACGG	421	QY	1320
QY	301	TCGTGAAATCTCTGTGTAGATCCTCTCAACATCCCGTACTATCCGAAATGCAAA	360	DB	1441
DB	422	TCGTGAAATCTCTGTGTAGATCCTCTCAACATCCCGTACTATCCGAAATGCAAA	481	QY	1380
QY	361	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTAT	420	DB	1501
DB	482	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTAT	541	QY	1440
QY	421	GGAGCATTTTCTGTATGATCGCTCCCAACAAATGCTACCGAGGCAAGTCTTC	480	DB	1561
DB	542	GGAGCATTTTCTGTATGATCGCTCCCAACAAATGCTACCGAGGCAAGTCTTC	601	QY	1500
QY	481	ACTGAAGGGAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTT	540	DB	1621
DB	602	ACTGAAGGGAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTT	661	QY	1560
QY	541	CGGCAAGGACAGGGTACCGTATATGATCTGACTTCTACTAATAATATTTGGACA	600	DB	1681
DB	662	AGGCAAGGACAGGGGTACCGTACATGATCTGACTTCTACTAATAATATTTGGACA	721	QY	1620
QY	601	AGTAACGGAACCAACAAATGACATGCTGATGTTTCGGGCTCTTCAAGAAATACAA	660	DB	1741
DB	722	ACAAATGGAACCAACAAATGACATGCTGATGTTTCGGGCTCTTCAAGAAATACAA	781	QY	1680
QY	661	ACAAAGAACCAACATGCTCGTCCAAAATACCTCCACCACTGCCCCACAGCCGTC	720	DB	1801
DB	782	ACGAAGAAATCAAAATGCTCGTCCAAAATACCTCCACCACTGCCCCACAGCCGTC	841	QY	1740
QY	721	GAGATCAAACTCACAAGCAACCACTGATGTCACCAACCACTCAATACCAAGCA	780	DB	1861
DB	842	GAGATCAAACTCACAAGCAACCACTGATGTCACCAACCACTCAATACCAAGCA	901	QY	1800
QY	781	AGTGTATGAGGACCTCGCAACATCCGCTCAGGCTCGGAGAACGAGAACCCACACA	840	DB	1921
DB	902	AATGATGAGGACCTCTAATACATCCGGTTCAGGGTCCGGAGAACGAGAACCCAT	961	QY	1860
QY	841	ACTTCTGATGCGGTTCACCAAGCAAGGGCTTTTCATCAACAAATGCCACCACTCC	900	DB	1981
DB	962	ACTTCTGATGCGGTTCACCAAGCAAGGGCTTTTCATCAACAAATGCCACCACTCC	1021	QY	1920
QY	901	CAACCAAGCAACGCAAGGAGGAGAAACAACAACCAATTCAGAGATGCTGTGACT	960	DB	2041
DB	1022	CAACCAAGCAACGCAAGGAGGAGAAACAACAACCAATTCAGAGATGCTGTGACT	1081	QY	1980
QY	961	GAACTAGACAAAATAACCAAACTGACCAAACTGCTCCATGCTCCCTCATTAACACT	1020	DB	2101
DB	1082	GAACTAGACAAAATAACCAAACTGACCAAACTGCTCCATGCTCCCTCATTAACACT	1141	QY	2040
QY	1021	ATCTCTACTAAACAACCTTCCAAACAACAACTTACGACTCTCTCTGCAACCAATTA	1080	DB	2161

MVIRPR	LOCUS	19112 bp	RNA	linear	VRL 22-JAN-1996
DEFINITION	Marburg virus (Popp)	NP, VP35, VP40, GP, VP30, VP24, L genes.			
ACCESSION	229337				
VERSION	229337.1	GI:450908			
KEYWORDS	GP protein; L protein; NP protein; vp24 protein; vp30 protein; vp35 protein; vp40 protein.				
SOURCE	Lake Victoria marburgvirus				
ORGANISM	Lake Victoria marburgvirus				
REFERENCE	1 (bases 2851 to 5818)				
AUTHORS	Bukreyev,A.A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.				
TITLE	The VP35 and VP40 proteins of filoviruses. Homology between Marburg and Ebola viruses				
JOURNAL	PEBS Lett. 322 (1), 41-46 (1993)				
MEDLINE	93245956				
PUBMED	8482365				
REFERENCE	2 (bases 5824 to 8669)				
AUTHORS	Bukreyev,A.A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.				
TITLE	The GP-protein of Marburg virus contains the region similar to the 'immunosuppressive domain' of oncogenic retrovirus P15E proteins				
JOURNAL	PEBS Lett. 323 (1-2), 183-187 (1993)				
MEDLINE	93265932				
PUBMED	8495737				
REFERENCE	3 (bases 1 to 19112)				
AUTHORS	Bukreyev,A.A., Volchkov,V.E., Blinov,V.M., Dryga,S.A. and Netesov,S.V.				
TITLE	The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain				
JOURNAL	Arch. Virol. 140 (9), 1589-1600 (1995)				
MEDLINE	96028047				
PUBMED	7487490				
REFERENCE	4 (bases 1 to 19112)				
AUTHORS	Bukreyev,A.A.				
TITLE	Direct Submission				
JOURNAL	Submitted (13-JAN-1994) Bukreyev A.A., Institute of Molecular Biology, Laboratory of molecular virology, Koltsovo, Novosibirsk region, Russia, 633159				
FEATURES	Location/Qualifiers				
source	1..19112				
	/organism="Lake Victoria marburgvirus"				
	/mol_type="genomic RNA"				
	/strain="Popp"				
	/db_xref="taxon:11269"				
	103..2190				
gene	/gene="NP"				
CDS	103..2190				
	/gene="NP"				
	/function="nucleoprotein"				
	/citation={3}				
	/number=1				
	/codon_start=1				
	/label=NP				
	/product="NP protein"				
	/protein_id="CAA82536.1"				
	/db_xref="GI:450909"				
	/db_xref="GOA:P35263"				
	/db_xref="UniProt/Swiss-Prot:P35263"				
	/translation="MDLSLELGTKTAPHVRNKKVILFDTNHQVSIQNCIIDAINS GIDLGLEGLTLCVHYNSDKDFNTSPIAKYLRDAGYDFVVKUADATRLDV IPNEHYSPILALAKTLESTESQGRIGLPLGCSLFLPKLVGDRASIEKALROVTV HQBQGVTVPNHWLTGTHMKVIFGILRSSFLFKVLHOGVNLVTGDYADYSI NSV GQTRFSGLLIVKTVLPIQKTDGVTLHPLVRTSKVKNVASFKQALSNLARGEYA PFARLNISGNLNEGLYQLPSAIGALVATGHATSGVAGNVEQYQQUREAAHDAIE KLQRHEHQBQIAIEDDEKTEQLEFHLQKTEITHSQTLAVLSKREKLAAAEIE NNTVEGQKQSQVQSFLNDPTPEVTYVQARPNRPTALPPVPVDNKIEHSTEDS SSSSSFDVLDNFDALLNEDDTLDDSVMLPTSTTSREFQGPAPPROSDLNNSQKOE DESTNPILKQFLYQLPLPVQEDSEYTTDSQESIDQPSDNEQGVDPPLPYAQE KRQDPTQHPAVSSQDFSGIDGVNDGILPIRSPSPSPAQEDTRAREAYELSPDFTN YEDNQQRVQPVVTKKGRFLYLVNDLLQTNPPESLITALVEYQNPVSAKELQADWFD MSFDERRHVAMNL"				
	2944..3933				
gene					

gene

```
/number=5
/codon_start=1
/label=VP30
/product="vp30 protein"
/protein_id="CA82540.1"
/db_xref="GI:450913"
/db_xref="GOA:P41326"
/db_xref="UniProt/Swiss-Prot:P41326"
/translation="MQPRGRSRTRNHTQASSIYHETQLPSKPHYTNHHPRRMSST
RSSAESPTNHIPRAPPPTNLSKPPPPKDMCRNMKIGLPDPTCNVRDHDLDNLT
NRELLLMARKWLPNTDFTFSLQDCGSPSLSKGLSKDQSQTKDVLTLNLHLNLY
LHRSIDGKDETSRLAASLTCAGIRKTRNRLINTWTEIHINHENLPQDQNGVINKQTY
TGIHLDKGGQFEALMQGWDKRSISLFVQAALYVNNNIPCESSTSVQASVDFHILPOS
QSKGQ"
gene
10205..10966
/genes="VP24"
CDS
10205..10966
/genes="VP24"
/citation={3}
/number=6
/codon_start=1
/label=VP24
/product="vp24 protein"
/protein_id="CA82541.1"
/db_xref="GI:450914"
/db_xref="UniProt/Swiss-Prot:P41325"
/translation="MAELSTRYNLPANVTEKSIINDLNSTARWIKEPSVGGWTVKWN
FVFHPTNGMALHHLKSNFVVPWQOTRNLPSHLFKNPKSTIIIPPLALRLLGLVAL
KDEELQOSLIQCFRSHVHMLSEMLLEVTSAIHISPNLLGIYLTDFMFKILMVGKVF
FNKMFILHVNDHGKPSIEIKLTQQIILITRVNMGFLVFEVRRIDIEPCGETVLSES
VVFGLVAEVAUREHSQMEKQGLDUTQYNSKIAI"
gene
11479..18474
/genes="L"
CDS
11479..18474
/genes="L"
/function="major component of RNA-dependent
RNA-polymerase"
/citation={3}
/number=7
/codon_start=1
/label=L
/product="L protein"
/protein_id="CA82542.1"
/db_xref="GI:450915"
/db_xref="GOA:P35262"
/db_xref="UniProt/Swiss-Prot:P35262"
/translation="MQHPTQYDARLSPIILDQCDLLARSLGLYSHYSHNPKLNR
IPHHYLRKNSLTALFTQNGSILTVPPHSHWDHILTSIQYDAINHVDVDFKLLPSEL
VKYAWNDEFKAYLNKILGLDHPFASARQWEDFSKPNPYWGMLLLVHLSQAR
RIKQGRGSLRNMKEITDLELFGIADFIIFKVPVKTIIRNAVSLOASKPGRLRVMYRD
QNLTPYLCDDDEFIVSASVECFIMIKDVFIERNTWEICARAWLEDSDGADYLPDLVL
GELYNQGDQI IAMYLEDGPKLKHLEPLCVSCIOTHGIFTGKWFQSFQRIESYVEEL
CSLNWKFILSGNKASCAQNFIKITIIQGLKTPQCYCELSLQKHGHWGHPVYIDVALDKV
KHAQSVKILKPKWFEFTFCVFKFIVAKNHHYSQGSWYKTTMDLHLTPLYLRQHVSN
FPSQAEIYQHLWEVFEHPELUFSTFKIISDLSIFIKDRATAVNQCWDSVDFRSVLGY
Query Match
Best Local Similarity 82.7%; Score 1692.4; DB 14; Length 19112;
Matches 1825; Conservative 0; Mismatches 221; Indels 0; Gaps 0;
QY 1 ATGAAGACCAATGTTTCCTTATCAGTCTTATCTTAATTTCAAGGGACAAAAAATCTCCCC 60
DB 5940 ATGAAGACCAATGTTTCCTTATCAGTCTTATCTTAATTTCAAGGGACAAAAAATCTCCCC 5999
QY 61 ATTTTAGAGATAGTAGTAATTAATCAACCCCAAAATGTGGATCGGTATGCTCCCGAAT 120
DB 6000 ATTTTAGAGATAGTAGTAATTAATCAACCCCAAAATGTGGATCGGTATGCTCCCGAAT 6059
QY 121 CTCGAGAACAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
DB 6060 CTCGAGAACAGAGATGTTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 6119
QY 181 GATTCCCCCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
```

```
DB 6120 GATTCCCCCTTTGGAGGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 6179
QY 241 GTTCAGTACACAGAGGGGGAGGAGCCAAAAACATGCTACAATATAGTGTACAGCATCCC 300
DB 6180 GTTGTGATATACAGAGGGGAGGAGCCAAAAACATGCTACAATATAGTGTACAGCATCCC 6239
QY 301 TCTGAAAAATCCTTGTCTTTAGATCCTCCTACCAACATCCGTGACTATCCGAAAAATGCAAA 360
DB 6240 TCTGAAAAATCCTTGTCTTTAGATCCTCCTACCAACATCCGTGACTATCCGAAAAATGCAAA 6299
QY 361 ACTATCCATCATATTAAGGTCAAAACCTCATGCAAGGGGATCGCCCTTCATTTATGG 420
DB 6300 ACTATCCATCATATTAAGGTCAAAACCTCATGCGCAAGGGATCGCCCTTCATTTGTGG 6359
QY 421 GGAGCATTTTTTCTGTATGATCGCATTCCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 480
DB 6360 GGAGCATTTTTTCTGTATGATCGCATTCCTCCCAACAAATGTACCGAGGCAAGGTCTTC 6419
QY 481 ACTGAAGGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTCTCG 540
DB 6420 ACTGAAGGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTCTCG 6479
QY 541 CGGCNAGGACAAAGGTACCGTACATATGATCTGACTTCTACTAATAAATATTTGGACAGT 600
DB 6480 AGGCAAGGACAAAGGTACCGTACATATGATCTGACTTCTACTAATAAATATTTGGACAGT 6539
QY 601 AGTAAACGGAAACGCAAAACAAATGACACTCGATGTTTTCGGCGCTCTTCAAGAATACAAATTC 660
DB 6540 AACATAGGACAAACAAACAAATGACACTCGATGTTTTCGGCGCTCTTCAAGAATACAAATTC 6599
QY 661 ACAAGAAACCAAAACATGTGCTCGTCCAAAATACCTCCACCACTGCCACACGCCCGTCCG 720
DB 6600 ACGAAGAAACCAAAACATGTGCTCGTCCAAAATACCTCCACCACTGCCACACGCCCGTCC 6659
QY 721 GAGATCAAACTCAACAGCACCCCAACTGATGGCACCACCACTCAATACACGAGACCCCAAGC 780
DB 6660 GAGATCAAACTCAACAGCACCCCAACTGATGGCACCACCACTCAATACACGAGACCCCAAG 6719
QY 781 AGTGATGATAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACACA 840
DB 6720 AATGATGATAGGACCTCATAACTCCGGTTCAGGGTCGGAGAACGAGAACCCCTATACA 6779
QY 841 ACTTCTGATGCGGTCAACAAAGCAAGGGCTTTTCATCAACAATGCCCCACCACTCCCTCAACA 900
DB 6780 ACTTCTGATGCGGTCAACAAAGCAAGGGCTTTTCATCAACAATGCCCCACCACTCCCTCAACA 6839
QY 901 CAACCAAGCACCCCAACAGCAAGGAGGAAACAAACAAACCACTTCCCAAGATGCTGTGACT 960
DB 6840 CAACCAAGCACCCCAACAGCAAGGAGGAAACAAACAAACCACTTCCCAAGATGCTGTGACT 6899
QY 961 GAACTAGACAAAAATAACACAACTGCACACCGTCCATGCCCCCTCATATAACACTTACCACA 1020
DB 6900 GAAACCAAAAAACCAACAAACGCGCACACCGTTCATGCCCCCCCCCAACCACTTACCACA 6959
QY 1021 ATCTCTACTAAACACACCTCCAAAACAACTTCAGCACTCTCTCTGCACTTATCAAAAAC 1080
DB 6960 ATCTCTACTAAACACACCTCCAAAACAACTTCAGCACTCTCTCTGCACTTATCAAAAAC 7019
QY 1081 ACCACCAATGACACACACAGAGCACTACTGAAAATGAGCAAAACGATGCCCCCTCG 1140
DB 7020 ACCACCAATGAGACACACAGAGCACTACTGAAAATGAGCAAAACGATGCCCCCTCG 7079
QY 1141 ATAACAAACCTCGCTCCAAACGAGGAAATCCCAACAGCAAGAGGACCAACGAGCAGCAAAAA 1200
DB 7080 AAAACAAACCTCGCTCCAAACGAGGAAATCTTACCAGCAAGAGGACCACTAACCAACAGAAA 7139
QY 1201 GGGCCCCGACCAACAGGACCCCAACACGACAAATGAGCACTTTTCCAGGCTCTCCCCCACC 1260
DB 7140 GGGCCCCCACCACCAACAGGACCCCAATATGACAAATGGGCATTTTAAACGAGTCTCCCCCACC 7199
QY 1261 CCCAGCTGAGTGCACAAACATCTTGTATATTTTCAAGAAAGAGGATCGATTTTCTGGAAA 1320
```


On Aug 27, 2003 this sequence version replaced gi:33868618.

```

COMMENT
FEATURES
    source
        location/Qualifiers
            1. .19151
                /organism="Lake Victoria marburgvirus"
                /mol_type="genomic RNA"
                /strain="M/S.Africa/Johannesburg/1975/Ozolin"
                /db_xref="taxon:11269"
                /note="crNA sequence"
            1. .19
            123. .2210
            /gene="NP"
            123. .2210
            /gene="NP"
            /note="nucleoprotein"
            /codon_start=1
            /product="NP protein"
            /protein_id="AAQ55255.1"
            /db_xref="GI:33868619"
            /translation="MDLHSLLELGTPTAPHVRNKKVILPDTTHQVSIQNIIDAISN
            GDIDGLGEGLLTLCVEHYNSDKDFTPIAKYLRDAGYDFVINKNADATRFIDLV
            INEHPHSLIILALKTLESTESQGRIGLFSCLPLKLVVGDRASTIEKALRQVTV
            HQEQGIVTPNHLATTGHMKVIFGLIRSSFILKFLVLIHOGVNLVTGHDADVDSIISNV
            GQTRPSGLLIIVNLEFLIQTDSGVTLHPLVRTSKVKEVASFKQALSNLABHGA
            PVARUNLGGINNLHGLYPQLSAYALGVATAGSTLGVNVGEQYQQLEAAHDAEVA
            KLQRHHEQIQAIAEDDERKILQEFHIQKTEITHSQTLAVLSQKREKLARLAAEIE
            NNIIVEDQGFQKSNRQFNDLPTPEVTVQARVNPRTALPPVDDDKIEHSTDS
            SSSSFDVLDNLPALINDEDITLDDSVMIPTSTSRFQGIPTSPGQSDLDNSQKQE
            DSTNPLIKQFURYQELPQVQDSEYITDQOESIDQPSDNEQGVDLPPPLPYLQE
            KRQDPIQHPAASQDPGSGIDGVNGDILFPIBSPSPSAQEDTRAREAYELSPDTFN
            YEDNQONQFQVVTKKGRFLFPNDLLQTSPPESLVTALVEEYQNVPVSAKELQADPFD
            MSFDERRHVAMNL"
            2364. .3953
            /gene="VP35"
            2364. .3953
            /gene="VP35"
            /note="putative minor component of RNA-dependent RNA-
            polymerase"
            /codon_start=1
            /product="VP35 protein"
            /protein_id="AAQ55256.1"
            /db_xref="GI:33868620"
            /translation="MWDSSIMQVSEGLMTGKVPIDQVFGANPLEKLYKRRKPKGTVG
            QCSPLCMSKATSTDIIWDQIVKTKLADLILPINRQSDIQSTLSEVTRVHEIER
            QLHEITPVLNMGRTGLEAIKSGHSEMLAKYDHLIVISGTRTTAPAAADAVLNHGVPMP
            QPAIKFDLGAQAQCSKGMVMAKTDAAADKMSKVLSEETFSKPNLSAKDLALLFL
            THLNGNPFHILAOVLSKIAVKSGKSGAFDLDAFHQILSEGANAQALFRLSTLDF
            LGVPPVIRVKNPQTVPRCQKSLRAVPNPNIDKGVVCVYSSEQGETRALKI"
            4587. .5498
            /gene="VP40"
            4587. .5498
            /gene="VP40"
            /note="putative matrix protein"
            /codon_start=1
            /product="VP40 protein"
            /protein_id="AAQ55257.1"
            /db_xref="GI:33868621"
            /translation="WASSNTNYTMOYLNPPPYADHGANQLIPADQLSNQHGITPNVV
            GDNLDDQFGNVCHTAELIIDSAYNERTVKGPVAPLPLGIMSNFEYPLAHTVAA
            LITGVTITQPTHNGQKQFVVRNLGTGIPAHFLRMLREGNAQFIQNNVIPRNFSTNQF
            TNNLNLVLSQKLDPAWRPSKDKLIGNTHMGPASIHENLPPIVLPTVKKQAYRQHK
            FNNNGPLLAISGLIHQLRYEKVPEKTSFLRFISLPADMFVYKMGMKRGESSPVVYFQ
            APENPLNGFNRRQVVLVANPNTLSAI"
            5960. .8005
            /gene="GP"
            5960. .8005
            /gene="GP"
            /note="envelope transmembrane glycoprotein"
            /codon_start=1
            /product="GP protein"
            /protein_id="AAQ55258.1"
            /db_xref="GI:33868622"
            /translation="MRTTCLFISLIIQKTLPILEIASNDQPNVDVSCSGTLQK
            EDVHLMGFTLSQKVADSPLEASKRWAFRTVPPKVEYTEGEAKTCYNISVTDSPSG
            
```



```

KSLLDPPTNVRDYPKCKTIIHHIQONPHAQIALHLWGAFPLVDRIASITWYRGKVF
TEGNTAAMIANKTVKHPHPIPSQGGYGRHNMNLTSTNKYWTSSNGTQNDTCGCTGHE
NESTNCAPKSTVPPPTAPRPEIKPTPTDTRLNTNPNSSDDEDLTSSGSGEQ
ELYTTSDAVTQGLSGLTTPSPQPPQNTNHSQDATTEDLNTNTTAQPPPT
SHNTTITGNTNLSKHLGTPAPQONTNPNQSMATENEKTSAPKALPTTENPTT
EKSTNNTKSPPTMBENTTNGHSTSSSTNSTOHLIYFRKRSILWEGDMFPPLDG
LINAPIDPVPNTKTIIDESSSGASTNEEQHASSNLSLTLVSLPCHSNTNATSGEN
ENDCAERLWSVQSDLAAGSLWIPFGPGLEGITTAGLTKNQNLVQNRNLANQT
AKSLDLVRVTTEERTFSLINRHAIIDFLLRWGTGCKVLGPDCCIGIEDLSRNTSEQI
DQIKKDEQKEGTGWLGGKMWTSDMGVLNTLNLGILLLLLSIAVLIALSCICRIFTYIG"
8888..9733
/gene="VP30"
CDS
8888..9733
/gene="VP30"
/codon_start=1
/product="VP30 protein"
/protein_id="AAQ55259.1"
/db_xref="GI:33868623"
/translation="MQPRGRSRTRNHQAIPSIYHETQLPSKPNVTNHHPRARMSST
RSSTESSPTNHIIPRAPPTSTFNLSKPPPPKMDMCRNMKIGLPCTDLTCNRDHDLDNLT
NRELLLMARKMLPNTDVFKSPQDCGSPSLSKGLSKDKQEQTKDVLTLNGLHILNY
LHRSBGLKDETSIAAALSLTCAGIRKTRNLSINTMTLHINHENLPQDQNGVTKQTY
TGIHLDKGQPEAALWQGDWRKESISLFVQAALYVNNNIPCESSISVQASDYHFLPQS
QKGCQ"
10226..10987
/gene="VP24"
CDS
10226..10987
/gene="VP24"
/codon_start=1
/product="VP24 protein"
/protein_id="AAQ55260.1"
/db_xref="GI:33868624"
/translation="NAELSTRVNLNPANVTEKSIINDLNSTARWIKEPSVGGWTVKMGV
FVFHPTMGMTLHLHKNFVFWPQWQTRNLFSLFKPNKSTIIEPFLALRILGLVAL
KQDELQSLPGFRSIVHMLSEWLLVETSAIHISPNLIGLYLTSDFMKILMAGVNF
FNKMTLHVANDHGRPSIEIKLTGQQLITRVNMGFLVEVRRIDIEPCCEGTILSES
VFGVLVAERKQHEKQMEKQPLNLTQYMNKIAI"
11500..18495
/gene="L"
CDS
11500..18495
/gene="L"
/notes="major component of RNA-dependent RNA-polymerase"
/codon_start=1
/product="L protein"
/protein_id="AAQ55261.1"
/db_xref="GI:33868625"
/translation="MQHPTQYDARLSSPVLVDQDLARSILGLYSHVSHNPKLRNCR
IPHHIYRLRNSTALTKTFLQNCISLTVPHSINWHLITSIQYDAINHVDVDFKYLPSL
VYANWDFELKTYLNLKILRDHVPFARSQCEDFSKPNPYTWGMGLLVHLSQLAR
RIKQGRSLRNMKPIGDLLEFGIADFIIFKVPVKTIIRNAVLSQASPKGLRVMYRD
QNLTPYLCDDPEIVSAVSECFIMIKDVFIERNTWEICARAWLEDSDGADYPLDVL
GELVYNGDOIYANYLEDGFKLTKHLEPLCVSCIQTHGFTPRKYWFQSQMIKSYDEL
CDNLKLARDSNKSCAQNFIKTIIOARLTPOCYCELSLOKMGHPVLYNDVALDKV
KHAQATKILKPKVMPFECVPKFIYAKNHHYSQGSWYKTHDLHLTPYLRLQHVSN
FPQAEIYQHLMEWTFVEHEPLFSTKISDLSIFIKDRATAVNQECWDSVDFRSVLGY
NPNVFRQSKRVEPQLQADSLNQLIDFAEKLEYLAPS YRNFSLKXELNLTGRF
GKLPYRVNVTQTLAALLADGLAKAPPSNMVVTTEREQEALLQWHHNSASITGEN
AIVRASVFTDLEKALLAFRYBFRHFNYCNRCYGVKNLPDMMHFLIPLCYMVHSDF
VSPHCVTEDNENPPDCANAVHYHLGGIEGLQOKLWTCISCAOITLVELTKLKLKS
SVMGNQCKITLUSLFPIDAPNDYQNEAELNARVABLAITGSGILPKPEETFVH
SGFIYFGKLVGQVQLPSLTKMARCPGLSDSIFDDLOGLSASIGTSFERGTSETRH
IPPSWFIASFHMLAINLNQHLFGFLFNIDISCFKPLTFSKLLAJITPOVLGG
LSFLNPEKFLVNSDPLTSGLFOLKNALEFLGGLFVILIAXKPKGLADAPVWNP
LGLNVPGSREIITFURQTVRENTISQNRINSLPHIGSDELDOKCEWLLSNPNVM
SRFADIPSRTPSGKRELQVLGVECTRTLLASRTISLTECTMLMKLRELTRNWKSW
PSYDALDDLSLESKEFTCTVDVANFLRAYSWSDVLGKRLIGATLPCLEQFKYKW
INLSDELREQFLGSESELITNLLPYDCKEURLGGSHDTELNYVSCALDRKVVQKHP
VRLWMTGNRAPYIGSRTEDEKIGYPLRVNCPSPAALEKAEIEMVSRLLAWVTQGTADRE
KLLIPLANSRVNLDQTVNLFLPTHYSGNIVHYNDQYQHSFMANRMSNTSTRAIIS
TNTLGYAGGQVADISNIIIFONTINLGVAVLDIALSLAKLSASNVTFRMLNKKCT
RHVPSEYLPFDKPLVDNLKNYMDNELVYNDPLCSGIKRGLGRVSRSTLSLNVSDI
GSYDPTTAAWTLGTITVGSIFSDSSSQSTDPISSGCTKTFVTHFLVYPVESIYFAPG
ANLIVESLSLRKISIKNLSIDTFLISSTIRNLSHRSDRILOSTIFRHLVLRLAHHI

```

```

PLISLMLGSGAGEKSSDAVRLLELTASVQNFHNFSCIMKKGOSLPWLIVFPSEGOQ
LKPILKILQRLSDLLSPDKVQRCRTKLTADTCCSIDSSVFWYPSKSTRTHYASLAWYRD
KANKVQPTFSLHLNCSFLESLSTSSVSNQVTSNKYIVHPDSKSTRTHYASLAWYRD
GSTALQRMIDKMPSEQLNVCNRPDKIRFKDNOKITKHQDQYKESQSQMSPE
NMOTFLAYMHNSSPSQITKSIDVHEDFASRVILNSKINNENLTDCTINTNLLTPTG
TEFLDTPLOSSSVYSTPRERSLLSREOASVLYVDCSNIPSIISLDPPGRNMSDONOIQ
MLINAYKEDLHACEDSNQCFPTGTVSSMYKLYDLPAPGELRAICLAEGESGARL
LLKWKETDHLFFNTLATDSQEAELISGRVIPRMLYINIDRLSALSLESRLINLNUITIQ
ITDITNPLMDSLVQIYLPEDSDILTDMAETTKTPYSSNAKNSEWLYLCCRRIRNPDKIS
IKVFLADYEGTFLFMKNAIQYQGVQKPKPYSSNAKNSEWLYLCCRRIRNPDKIS
DOVGFLITCAMSQRQAIPLYKHIENKYPASLHPEFLTLGFLPSLSSFSFCHRYTIPF
SEGKALFHKQSVYVRQKQHLHLSMLYENNSPLDLRNHPIFCSLRGKITKYVNDILK
LNLVIKAVEKGKWSQVETLPNNHVSVCIVHVDHECTGCEKEXBLLLKLLDFIRNTRIAEQ
KLMNRVIGIILFPFGLSKRSRRA"
complement (19133..19151)
ORIGIN
Query Match 80.8%; Score 1653.6; DB 14; Length 19151;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 1800; Conservative 0; Mismatches 244; Indels 0; Gaps 0;
Qy 1 ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACACAAAAATCTCCCC 60
Db 5960 ATGAGGACTACATGCCCTTTATCAGTCTTATCTTAATTCAGGGATATAAAACTCTCCCT 6019
Qy 61 ATTTAGAGATAGTAGTAATAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAATCT 120
Db 6020 ATTTGGAGATAGGACGATCAACCCCAAAATGTGGATTGGTATGCTCCGGAATCT 6079
Qy 121 CTCAGAGACAGAGAGCGTCCATCTGATGGGATTCACACTGAGTGCGGCAAAAATGCTCT 180
Db 6080 CTCAGAAAAACAGAAAGATGTCATCTGATGGGATTTACACTGAGGGGCGAGAAAGTGTCT 6139
Qy 181 GATTCCCTTTGGAGGCATCAAGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 240
Db 6140 GATTCCCTTTGGAGGCATCAAGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 6199
Qy 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTCAATATAAAGTGTAAAGCGATCCC 300
Db 6200 GTTGAGTACAGGAGGAGGAGGAGCCAAACATGCTCAATATAAAGTGTAAAGCGATCCC 6259
Qy 301 TCTGGAATATCTTGTCTTTAGATCTCTACCAACATCCCTGACTATCCGAAATGCAAA 360
Db 6260 TCTGGAATATCTTGTCTTTAGATCTCTCCACCAACGTCCTGACTATCTTAAATGCAAA 6319
Qy 361 ACTATCATCATATTCAGAGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTATGG 420
Db 6320 ACTATCATCATATTCAGAGTCAAAACCCCTCATGCAGGGGATCGCCCTTCATTATGG 6379
Qy 421 GGAGCATTTTTTCTGTATGATCGCATTCCTCCACAAATGTATCCGAGGCAAAAGTCTTC 480
Db 6380 GGAGCATTTTTTCTGTATGATCGCATTCCTCCACAAATGTATCCGAGGCAAAAGTCTTC 6439
Qy 481 ACTGAAGGACATAGCAGCTATGATGTCAATAGCAGTGCACAAATGATTTCTTCG 540
Db 6440 ACTGAAGGACATAGCAGCTATGATGTCAATAGCAGTGCACAAATGATTTCTTCG 6499
Qy 541 CGGCAAGGACAAAGGTCACCGTCAATAGTCTGACTTCTACTATAAATATTTGGAACAAGT 600
Db 6500 AGGCAGGACAAAGGTCACCGTCAATAGTCTGACTTCTACTATAAATATTTGGAACAAGT 6559
Qy 601 AGTAAACGGAACGCAAAAGAAATGACATGATGTTTTCGCGCTCTTCAAGAATACAAATCT 660
Db 6560 AGCAACGGAACGCAAAAGAAATGACATGATGTTTTCGCGCTCTTCAAGAATACAAATCT 6619
Qy 661 ACAAGAACCAAAATGTGCTCCGTCGCAAAATACCTCCACCACTCCGCCACACCGCTCG 720
Db 6620 ACGAAGAACCAAAATGTGCTCCGTCGCAAAATACCTCCACCACTCCGCCACACCGCTCG 6679
Qy 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780
Db 6680 GAAATCAAAACCAACGACACCCCAACCGATGCCACTAGACTCAACACCAACCAACCAAC 6739

```

QY 781 AGTGATGATGAGGACCTCGCACATCCGCTCAGGGTCGGGACGAGGACGAGACCCACACA 840
DB 6740 AGTGATGATGAGGATCTCACAACTCCGGCTCGGGGTCTGGGAAACAGAGAACTCTATACG 6799
QY 841 ACTTCTGATGCGGTCAACAAAGCAAGGGCTTTCATCAACAATGCCACCCACCTCCCTCACCA 900
DB 6800 ACTTCTGATGCGGTCACTAAGCAAGGGCTTTCATCAACAATGCCACCCACCTCCCTCACCG 6859
QY 901 CAACCAAGACGCCACAGCAAGGAGGAAACAACAACCAATTCCTCAAGATGCTGTGACT 960
DB 6860 CAACCAAGGACGCCACAGCAAGGAGGAAACAACAACCAATTCCTCAAGATGCTGTGACT 6919
QY 961 GAACTAGACAAATAACACAACTCCACAAACCGTCCATGCCCTCCATCAACACCACTACCA 1020
DB 6920 GAACTTGATAACCAACAACAACTCCACAAACCGCTCCATCAACACCACTACCA 6979
QY 1021 ATCTCTACTAAACAACACCTCCAAAACAACCTTCAGCACTCTCTCTGCAACCAATTAACAAC 1080
DB 6980 ATCTCCACCAACAACCTCCAAAACAACCTTCAGCACTCTCTCTGCAACCAATTAACAAC 7039
QY 1081 ACCACAATGACAACAACAGCAACATCACTGAAAAATGAGCAAAACAGTGCCCTCCG 1140
DB 7040 ACCACAATCCCAACAACAAGCATGGCCACTGAAAAATGAGAAAAACAGTGCCCTCCG 7099
QY 1141 ATAAACAACCTGCTCCAAACGGGAATCCACCAACAGCAAGAGCAGCAGCAAAAA 1200
DB 7100 AAAGCAACCTGCTCCCAACAGAAAATCTTACCAACAGAAAAGAGCAACAATAACAAA 7159
QY 1201 GCGCCCGCACAAACGCGCAACCAACAGCAATGAGCAATTCACCAAGTCCCTCCAC 1260
DB 7160 AGCCCAACCAATGCAACCAATGCAACCAATGCAATTCACCAAGTCCCTCCAC 7219
QY 1261 CCCAGCTGAGTGCAACAACCTCTGATATTTTCAAGAAAGAGCAAGCAATTCCTCGAAA 1320
DB 7220 CCCAACTGCAACAACTCTTATATATTTTCAAGAAAGAGCAAGCAATTCCTCGAGG 7279
QY 1321 GAAGTGATATTTCCGTTTTCAGTGGTAAATAATGCAATTAATGATTTGATCCA 1380
DB 7280 GAAGGAGCATGTTCCCTTTTTCAGTGGTAAATAATGCAATTAATGATTTGATCCA 7339
QY 1381 ATCCCAACACAGAAACAACTTTGATGAATCTCCAGCTTTAACTTCAACTAATGAG 1440
DB 7340 GTTCCAAATACAAAGCAACTTTGATGAATCTTCTAGTTCTGGTCTCACTGAGAA 7399
QY 1441 GAACCAACACATCCCGCAATATCAGTTTAACTTTCTCTTATTTTCTGATAAAATGGA 1500
DB 7400 GATCAACATGATCTCCCAATATCAGTTTAACTTTCTCTTATTTTCTTCAACAGTGA 7459
QY 1501 GATCTGCTACTCTGGGGAACAGCAATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 7460 AACACTGCTACTCTGGGGAATGAGATGATGATGATGATGATGATGATGATGATGATGAT 7519
QY 1561 GTGAGGAGGAGATTTGGCGGAGGCTTACGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 7520 GTTCAAGAGGAGGATTTGGCGGAGGCTTACGATGATGATGATGATGATGATGATGATGATGAT 7579
QY 1621 GAAGGACTCTATCTGCGGTTTAAATCAAAAATCAGAAATTTAGTTTGTAGTTGAGG 1680
DB 7580 GAAGGACTTTATCCGACAGTTTAAATCAAAAATCAGAAATTTAGTTTGTAGTTGAGG 7639
QY 1681 CGTTAGCTAATCAACTGCTAAATCTCTGGAGCTCTTGTAAAGGCTCAACACCGAGAA 1740
DB 7640 CGTCTAGCAATCAAACTGCAAAATCTTGGAACTTTACTAAGGCTCAACACCGAGAA 7699
QY 1741 AGGACATTTTCTTAAATCAATAGGATGCAATGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
DB 7700 AGAACAATTTTCTTAAATCAATAGGATGCAATGATTTTCTTCTTCTTCTTCTTCTTCTTCT 7759
QY 1801 ACATGCAAGGCTCTAGGACTGATTTGCTGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 7760 ACATGCAAGGCTCTAGGACTGATTTGCTGATGATGATGATGATGATGATGATGATGATGAT 7819
QY 1861 TCAGAACAAATCGAACAAATCGAAGGATGAACAAAGGAGGAACTGGCTGGGGTCTA 1920

DB 7820 TCAGAACAAATCGAACAAATCGAAGGATGAACAAAGGAGGATGGTTGGGGCTCG 7879
QY 1921 GGTGGCAAAATGGTGACATCTGACTGGGGTGTCTTCAACAAATTTGGGATCTCTCTACTA 1980
DB 7880 GGTGGTAAATGGTGACATCCGACTGGGGTGTCTTCAACAAATTTGGGATCTCTCTACTA 7939
QY 1981 TTATCTATAGCTGTTCTGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2040
DB 7940 TTGTCATAGCTGTTCTGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7999
QY 2041 GGAT 2044
DB 8000 GGAT 8003

RESULT 11
AF005733
LOCUS
DEFINITION Marburg virus strain M/S.Africa/Johannesburg/1975/Ozolin
glycoprotein precursor (GP) gene, complete cds.
ACCESSION AF005733
VERSION AF005733.1 GI:2459875
KEYWORDS
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 2046)
AUTHORS Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and
Feldmann,H.
TITLE Variation in the glycoprotein and VP35 genes of Marburg virus
strains
JOURNAL Virology 240 (1), 138-146 (1998)
MEDLINE 98110148
PubMed 9448698
REFERENCE 2 (bases 1 to 2046)
AUTHORS Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
TITLE Submitted (29-MAY-1997) Special Pathogens Branch, Centers for
Disease Control and Prevention, 1600 Clifton Road, Building 15,
Room SB611, Mail Stop G14, Atlanta, GA 30333, USA
FEATURES
source
1. .2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/S.Africa/Johannesburg/1975/Ozolin"
/db_xref="taxon:11269"
gene
1. .2046
/gene="Gp"
CDS
1. .2046
/gene="Gp"
/note="virion surface glycoprotein"
/codon_start=1
/product="glycoprotein precursor"
/protein_id="AAC40458.1"
/db_xref="GI:2459876"
sig_peptide
1. .81
/gene="Gp"
mat_peptide
82. .2043
/gene="Gp"

misc_feature		/product="glycoprotein" 1930..2019 /gene="gp" /note="encodes transmembrane anchor peptide"	
ORIGIN			
Query Match		80.8%; Score 1652.4; DB 14; Length 2046;	
Best Local Similarity		88.0%; Pred. No. 0;	
Matches 1800; Conservative		0; Mismatches 246; Indels 0; Gaps 0;	
QY	1	ATGAAGACACATGTTTCTCTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCC	60
DB	1	ATGAGGACTACATGCTCTTTATAGTCTTATCTTAATTCAGGGATAAAAATCTCC	60
QY	61	ATTTTAGAGATAGTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCG	120
DB	61	ATTTTGGAGATAGCCAGTAACGATCAACCCCAAAATGTGGATTCGGTATGCTCCG	120
QY	121	CTCCAGAGACAGAGACGTCCATCTGATGGATTCACACTGAGTGGGCAAAAGTTG	180
DB	121	CTCCAGAAAACAGAAAGATGTCATCTGATGGGATTTACACTGAGCGGGCAGAAA	180
QY	181	GATTCCTCTTGGAGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCA	240
DB	181	GATTCCTCTTGGAGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCA	240
QY	241	GTGTAGTACACAGAGGGGAGGAAGCCAAAACATGCTACAATATAAGTGAACG	300
DB	241	GTGTAGTACACAGAGGGGAGGAAGCCAAAACATGCTACAATATAAGTGAACG	300
QY	301	TCGTGAAATCTTGTGTAGATCTCTCTCAACATCCGTGACTATCCGAAATGCA	360
DB	301	TCGTGAAATCTTGTGTAGATCTCTCTCAACATCCGTGACTATCCGAAATGCA	360
QY	361	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCTTCATTTAG	420
DB	361	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCTTCATTTAG	420
QY	421	GGAGCATTTTCTGTATGATCGATTCCTCCACAACTGCTACCGAGGCAAGTCT	480
DB	421	GGAGCATTTTCTGTATGATCGATTCCTCCACAACTGCTACCGAGGCAAGTCT	480
QY	481	ACTGAAGGAAACATAGCAGCTATGATGTCATTAAGACAGTGCAAAAATGATT	540
DB	481	ACTGAAGGAAACATAGCAGCTATGATGTCATTAAGACAGTGCAAAAATGATT	540
QY	541	CGGCAAGGACAGGTACCGTATATGATCTGATCTTCTACTAATAATATTGGCA	600
DB	541	AGGCAAGGACAGGTACCGTACATGAACTGACTTCTACTAATAATATTGGCA	600
QY	601	AGTAACGGAAACGAAACGATGACACTGGATGTTTCGGCGCTCTTCAAGAAAT	660
DB	601	AGCAACGGAAACGAAACGATGACACTGGATGTTTCGGCGCTCTTCAAGAAAT	660
QY	661	ACAAAGAACCAACATGTGCTCCGTCCAAAATACTTCAACCACTGCCACAGCC	720
DB	661	ACGAAGAACCAACATGTGCTCCGTCTAAACACACCCCAACCCGACAGCCCG	720
QY	721	GAGATCAAACTCACAGCAACCCCACTGATGCGCCCAACCACTCAATACCA	780
DB	721	GAATCAAAACCAAGCAACCCCACTGATGCGCCCAACCACTCAATACCA	780
QY	781	AGTGTATGAGGACCTCCGCAATCCGGCTCAGGGTCGGGAGACGAGAACCC	840
DB	781	AGTGTATGAGGATCTCAACATCCGGCTCGGGTCTGGGGACAGGAATCTAT	840
QY	841	ACTTCTGATGCGGTCAACGAAGAGGCTTTTCATCAACAATGCCACCACTCC	900
DB	841	ACTTCTGATGCGGTCACTAAGCAAGGCTTTTCATCAACAATGCCACCACTCC	900
QY	901	CAACCAAGCAACCCCAAGGAGGAAACACACAAACCAATCCCAAGATGCTG	960
DB	901	CAACCAAGGCAACCCCAAGGAGGAAACACACAAACCACTCCCAAGATGCTG	960
QY	961	GAACCTAGACAAAAATAACACAACTGCAACACCGCTCCATGCCCCCTCAT	1020
DB	961	GAACCTGATAACACCAACACAACTGCAACACCGCCCGCTCCCAACACCA	1020
QY	1021	ATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTGCAACCA	1080
DB	1021	ATCTCCACCAACAACTCCAAACACAACTTCAGCACTCTCTGCAACCA	1080
QY	1081	ACCACCAATGACAAACACACAGAGCAATCACTGAAATGAGAAACCACTG	1140
DB	1081	ACCAACCAATCCCAACACACAAAGCATGGTCACTGAAATGAGAAACCACTG	1140
QY	1141	ATAACAAACCTCTCCAAACGGAATCCCAACAGCAAAAGAGCACAGCAGCA	1200
DB	1141	AAAGCAACCTCTCCAAACGGAATCCCAACAGCAAAAGAGCACCAACA	1200
QY	1201	GGCCCGCCACAAACGAGCAACCAACAGCAATGAGCAATTCACAGTCCCT	1260
DB	1201	AGCCCCACCAATTTGGAACCAATTAACAATGGAATTTCACTAGTCCCT	1260
QY	1261	CCCAGCTCGAGTCACACACATCTTTGTATATTTTCAAGAAAGCGATCG	1320
DB	1261	CCCAGCTCGACTACACACATCTTTGTATATTTTCAAGAAAGCGAGTAT	1320
QY	1321	GAAGGTGATATATTTCCCGTTTTTAGATGGGTTAATAATCTGAAATTTG	1380
DB	1321	GAAGGCGACATGTTTCCCTTTTCTAGATGGGTTAATAATGCTCCAATTT	1380
QY	1381	ATCCAAACACAGAAACATCTTTGATGAACTCCAGCTTTAATACTTCA	1440
DB	1381	GTTCCAAATACAAAGACAATCTTTGATGAACTTTCTAGTCTGGTGT	1440
QY	1441	GAAACACACACTCCCGCAATATCAGTTTAACTTTCTTATTTTCTGATA	1500
DB	1441	GATCAACATGATCTCTCAATATCAGTTTAACTTTAICTTATCTTCTCA	1500
QY	1501	GATCTGCTACTCTGGGGAAAAACGAGAAATGATTGTGATCGAGAGTTG	1560
DB	1501	AACACTGCTACTCTGGAGAAAAATGAGAAATGATTGTGATCGAGCTA	1560
QY	1561	GTGAGGAGGACGATTTGGCGGAGGGCTTAGCTGGATACCAATTTTGG	1620
DB	1561	GTTTCAGGAGGACGACTGGCAGAGGGCTCAGTTGGATACCACTTTT	1620
QY	1621	GAAGCACTTATCTGCGGTTTAAATCAAAAATCAGAACAATTTAGTT	1680
DB	1621	GAAGCACTTATACCGCAGGTTTAAATCAAAAATTTGGTCTCGAGTT	1680
QY	1681	CGCTTAGCTAATCAAACTGTCTAAATCTTTGGAGCTCTTTGTTAAG	1740
DB	1681	CGTCTAGCCAATCAAACTGTCAAAATCTTTGGAACTCTTACTAAG	1740
QY	1741	AGGACATTTTCTTAATCAATAGCATGCAATGCACTTTTTCCTTAC	1800
DB	1741	AGAACATTTTCTTAATCAATAGCATGCTATGCACTTTCTACTCAC	1800
QY	1801	ACATGCAAGGTGCTAGGACCTGATTTGTTGCTAGGAATAGAAATCT	1860
DB	1801	ACATGCAAGGTGCTTGGACCTGATTTGTTGCTAGGAATAGAAATCT	1860
QY	1861	TCAGAACAAATTCGAAATCAGAAAGGATGAACAAAAAGGAGGAACT	1920
DB	1861	TCAGAACAGATTGACCAATCAAGAAAGGACGAAACAAAAAGGAGG	1920
QY	1921	GGTGCAAAATGGTGACATCTGACTGGGTTGTTCAACCAATTTGGG	1980
DB	1921	GGTGTAATGGTGACATCCGACTGGGGTTTCTTACTAACTTTGGG	1980
QY	1981	TTATCTATAGTGTCTGATTTGCTCTGCTGTATCTTCTGATCTTCA	2040
DB	1981	TTGTCCATAGTGTCTGATTTGCTCTATCTTCTGATTTTGTCTTACT	2040

```
QY 2041 GGATGA 2046
Db 2041 GGATAA 2046

RESULT 12
LOCUS AX717725 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 13 from Patent W002079239.
ACCESSION AX717725
VERSION AX717725.1 GI:29890731
KEYWORDS
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
VIRUSES; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
1
REFERENCE Grogan,C.C., Hevey,M.C. and Schmaljohn,A.L.
AUTHORS Chimeric filovirus glycoprotein
TITLE Patent: WO 02079239-A 13 10-OCT-2002;
JOURNAL U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
    Location/Qualifiers
        1..2046
            /organism="Lake Victoria marburgvirus"
            /mol_type="unassigned DNA"
            /strain="Raven"
            /db_xref="taxon:11269"
            /note="chimeric molecule between Marburg virus
            Glycoprotein 1 and Marburg virus Glycoprotein 2"

ORIGIN
Query Match 74.2%; . Score 1518; DB 6; Length 2046;
Best Local Similarity 83.9%; Pred. No. 0;
Matches 1716; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

QY 1 ATGAAGACACATGTTCTCTATCAGTCTTATCTTAAATCAAGGACAAAAATCTCCCC 60
Db 1 ATGAAGACATATATTTCTGATGATGTTCTGATGATGTTCTGATGATGTTCTGATG 60

QY 61 ATTTTAGAGTAGCTAGTAATTAATCAACCCCAAAATGCGATGTTGCGTATGCTCCGGA 120
Db 61 GTTTTAGAAATGCTAGTAACAGCCCACTCAAGATGTAGATTCACTGCTGCTCCGGA 120

QY 121 CTCGAGAGACAGAAAGCTCCTCTGATGGATTTCACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCGAGAGACAGAAAGCTTTCATCTGATGGATTTCACTGAGTGGGCAAAAAGTTGCT 180

QY 181 GATTCCTCTTTGGAGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db 181 GATTCCTCTTTGGAGCATCTAAAGCATGGGCTTTTCAGGACAGGTGTACCTCCCAAG 240

QY 241 GTTGAGTACAGAGGGGAGAGAGCCAAACATGCTACAAATATAAGTGTAAACGGATCCC 300
Db 241 GTTGAGTACAGAGAGGAGAGAGCCAAACATGCTACAAATATAAGTGTAAACGGATCCC 300

QY 301 TCTGAAAATCCTGCTGTAGATCCTCTACCAACATCCGAGTATCCGAAATGCAAA 360
Db 301 TCTGAAAATCCTGCTGTAGATCCTCTACCAACATCCGAGTATCCGAAATGCAAA 360

QY 361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTGTTTCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420

QY 421 GGAGCATTTTCTGTATGATGCGATTCCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 480
Db 421 GGGCATTTTCTGTATGATGCGATTCCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 480

QY 481 ACTGAAGGAAACATAGCAGCTATGATGTTCAATAGACAGTGCACAAATGATTTTCG 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATGTTCAATAGACAGTGCACAAATGATTTTCT 540

QY 541 CGGCAGGACAGGGTACCGTCATATGAATCTGACTTCTACTAATAATATTTGGACAAGT 600
Db 541 CGGCAGGACAGGGTACCGTCATATGAATCTGACTTCTACTAATAATATTTGGACAAGT 600

541 AGCAAGGACAAAGGTTATCGTCATCATGAACTTGACCTCCACCAATAAATATTGGACAAGC 600
601 AGTAAACGGAACGCAAAACGAATGACATCGGATGTTTCGGCGCTCTTCAAGAAATACAATTC 660
601 AGCAATGAAACGACAGAGAAATGATACGGGATGTTTGGCATCTCTCCCAAGAAATACAATCC 660
661 ACAAGAACCAAAATGCTGCTCCGTCCAAATATCTCCACCATCTCCACCAAGCCGCTCG 720
661 ACAAAACATCAAAACATCCCTCCATCTCTTAAACCTCCATCCCTGCGCACAGTAACCTCG 720
721 GAGATCAAACTCAACAGCAGCCCACTGATGCCCAACAACTCAATACACAGGACCCCAAGC 780
721 AGCATTCACTACTACAAATCTCAAAATTAATCTGCTAAATCTGGAATATGAAACCAAGT 780
781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACCGAGAAACCCACACA 840
781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
841 ACTTCTGATGGGTACCAAGCAAGGGCTTTCATCAACAAATGCCACCACTCCCTCAACA 900
841 ACTCTTAATGTAGTCACTGAACAGAAACAAATCGTCAACAAATATTTGCCACTCTCTCACTA 900
901 CAACCAAGCAGCCCAACAGCAAGGAGGAAACAAACAAACCAATTCCTCCAGATGCTGTGACT 960
901 CATCAAGCAGCTCAACACATGACAAACAGTAGAATCTCTCCCGACATGCTGTAACT 960
961 GAACAGCAAAATAACAACTGCAACACCGTCCATGCGCTCCCTCATATAACACTACCACA 1020
961 GAGCAATGGAACCGACCCCAACAAACCAACACCAAGCAACGCTCTCAACAAATACTAATA 1020
1021 ATCTCTACTAACACACTCTCCAAACAACTTTCAGCACTCTCTCTGACCACTTACAAAC 1080
1021 ACTCCCACTTAAACACTCTCAAGTACAACTCTAGTACTCTCTCCCTCCCAACCGCAAC 1080
1081 ACCACCAATGACAAACACAGACGACAACTCACTGAAATGAGCAAAACGAGTCCCTCCG 1140
1081 ATCAACCAATATGATACACACAGTGAATCTAGCAGAAAGCAACAAACAAATGCTCAGTTG 1140
1141 ATAAACACCTCTCCACCGGAAATCCCAACACAGCAAGAGACCAAGAGCAAGCAAGCAAAA 1200
1141 AACCAACTCTAGATCCAAACAGAAAAATCCCAACACAGCAAGAGCAAGCAAGCAAC 1200
1201 GGGCCCGCAACAGGCAACCAACACAGCAAAATGAGCAATTCACGATCTCTCCCGCAC 1260
1201 AACATCATATGACGACATCAGATATAACAGCAACCAACCAACCAAAATCTCTCCGAT 1260
1261 CCCAGCTCGACTGCAACAACTCTTGTATATTTTCAGAAAGAAAGCGATCGATTTCTGGA 1320
1261 TCTAGTCGACAAACCCCGCTCTCTATATCTTTAGAAAGAAACGATCGATTTCTGGA 1320
1321 GAAGGTGATATATTCCTGTTTATAGATGGGTAAATAAATCTGAAATGATTTTGTATCCA 1380
1321 GAAGGTGATATATTCCTGTTTATAGATGGGTAAATAAATCTGAAATGATTTTGTATCCA 1380
1381 ATCCCAACACAGAAACAACTCTTGTATGATGATCTCCAGCTTTAATCTTCAACTAATG 1440
1381 ATCCCAACACAGAAACAACTCTTGTATGATGATCTCCAGCTTTAATCTTCAACTAATG 1440
1441 GAACACACACTCCCGCAATATCAGTTTAACTTTCTTATTTTCTTATTTTCTGATAAAATG 1500
1441 GAACACACACTCCCGCAATATCAGTTTAACTTTCTTATTTTCTGATAAAATG 1500
1501 GATATCTGCTACTCTCGGGGAAACGAGAAATGATTTGATGATGATGATGATGATGATGATG 1560
1501 GATATCTGCTACTCTCGGGGAAACGAGAAATGATTTGATGATGATGATGATGATGATGATG 1560
1561 GTGAGAGAGACGATTTTGGCGGACGAGGCTTAGCTGGATACCAATTTTGGCCCTGGAAATC 1620
1561 GTGAGAGAGACGATTTTGGCGGACGAGGCTTAGCTGGATACCAATTTTGGCCCTGGAAATC 1620
1621 GAGGACTCTATCTGCGGTTTAAATCAAAATTCAGAACAAATTTAGTTTGTAGCTTGG 1680
1621 GAGGACTCTATCTGCGGTTTAAATCAAAATTCAGAACAAATTTAGTTTGTAGCTTGG 1680
```

QY 1681 CCCTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTGTTAAGGGTCACAAACGAGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTGTTAAGGGTCACAAACGAGAA 1740
QY 1741 AGGACATTTTCCTTAATCAATAGGATGCAATTTGCTTACGAGTGGGGCGGA 1800
Db 1741 AGGACATTTTCCTTAATCAATAGGATGCAATTTGCTTACGAGTGGGGCGGA 1800
QY 1801 ACATGCAAGTCTAGGACCTGATTTGCTAGGATAGAGATCTATCTAAAAATATC 1860
Db 1801 ACATGCAAGTCTAGGACCTGATTTGCTAGGATAGAGATCTATCTAAAAATATC 1860
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGACAAAGGAGGAAACTGGCTGGGCTCTA 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGGATGACAAAGGAGGAAACTGGCTGGGCTCTA 1920
QY 1921 GGTGCAATGTGACATCTGACTGGGTGTTCTCACCATTGGGCATCTGCTACTA 1980
Db 1921 GGTGCAATGTGACATCTGACTGGGTGTTCTCACCATTGGGCATCTGCTACTA 1980
QY 1981 TTATCTATAGCTGTTCTGATTCCTCTGCTCTATCTGCTGATCTTCACTAAATACATT 2040
Db 1981 TTATCTATAGCTGTTCTGATTCCTCTGCTCTATCTGCTGATCTTCACTAAATACATT 2040
QY 2041 GGATGA 2046
Db 2041 GGATGA 2046
RESULT 13
AF005734
LOCUS
DEFINITION Marburg virus strain M/Kenya/Kitum Cave/1987/Ravn glycoprotein precursor (Gp) gene, complete cds.
ACCESSION AF005734
VERSION AF005734.1 GI:2459877
KEYWORDS Lake Victoria marburgvirus
SOURCE Lake Victoria marburgvirus
ORGANISM Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 2046)
AUTHORS Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and Feldmann,H.
TITLE Variation in the glycoprotein and VP35 genes of Marburg virus strains
JOURNAL Virology 240 (1), 138-146 (1998)
MEDLINE 98110148
PUBMED 9448698
REFERENCE 2 (bases 1 to 2046)
AUTHORS Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
TITLE Direct Submission
SUBMITTED (29-MAY-1997) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Building 15, Room SB611, Mail Stop G14, Atlanta, GA 30333, USA
FEATURES
Location/Qualifiers
1..2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/Kenya/Kitum Cave/1987/Ravn"
/db_xref="taxon:11269"
1..2046
/gene="Gp"
1..2046
/gene="Gp"
/note="virion surface glycoprotein"
/codon_start=1
/product="glycoprotein precursor"
/protein_id="AAC40459.1"
/db_xref="GI:2459878"
/translation="MKTIYFLISLILIQSIKTLPLVLEIASNSQPDVSVCSGTQKLT
EDVHLMGFTLSQKRVADSPLEASKRWAFRTGVPPKNVYETGEAEKTCYNISVTPDSG

KSLLDDPPSNIRDYPKCTVHHIOQNPHAGGIALHLWGAFFLYDRVASTMYRGKVF
TEGNIAAMI VNKTVHRMIFRGQGYRHMLNTSTNKYWTSSNETORNDTGCGILQEY
NSTNQCPSPSLKPSLPVTPTPSIHSTNTAKSGTMNPSDESDLMISGSGEQ
GPHTTLVNVTQKQSSTILSPHSSTQHEQNSTPNSHRHVAHTDPTTQFATL
LNNNTPTNTLYKNSLSTPPTNTTNDTQRELAESEOTNAQLNLTLDPTBNPTT
GQDNTSNLIIMTSDITSKHPNTSSPDSPTTPPLYFRKRSI FWKEGDI PFPDGG
LINTEDFDPIPNTEITFDESPSNTSNEQHPNLSLTFSPFKNGDGTASGEN
ENDCDALRWVQEDDLAAGLSWIPFPGIEGLYTAGLKNQNNLVCRRRLANQT
AKSLELLRVTEERTFSLINRAIDLTFRWGCTKVLGPDCCIGIBDLKNSISEQI
DKIRKQKQKETGMLGKQWTSDWGLVTNLGILLLLSIAVLIALSCIRFTKYIG"
1..81
/gene="Gp"
82..2043
/gene="Gp"
/product="glycoprotein"
1930..2019
/gene="Gp"
/note="encodes transmembrane anchor peptide"
sig_peptide
mat_peptide
misc_feature
ORIGIN
Query Match 74.0%; Score 1513.2; DB 14; Length 2046;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 1713; Conservative 0; Mismatches 333; Indels 0; Gaps 0;
QY 1 ATGAAGACCAATGTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAAATCTCCCC 60
Db 1 ATGAAGACCAATATTTCTGATTAGTCTCAATTTTAATCCAAAGTATAAAAAATCTCCCT 60
QY 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAATGTGGATTTCGGTATGTCCTCCGAACT 120
Db 61 GTTTTAGAAATGTCTAGTAACAGCCAACTCAAGATGTAGATTCTAGTGTCTCCGGAACC 120
QY 121 CTCAGAGACAGAGACGCTCCATCTGATGGGATTCACATCTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCAGAGACAGAGATGTTTCACTCTGATGGGATTCACATCTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTGGAGGCATCCAAAGGATGGCTTTTCAGGACAGGTGTACTCTCCCAAGAT 240
Db 181 GATTCCCTTTGGAGGCATCTAAACGATGGCTTTTCAGGACAGGTGTCTCTCCCAAGAAC 240
QY 241 GTTCAGATACAGAGGGGGAGAGCCAAACATGCTACAATATAAGTGTACGGATCCC 300
Db 241 GTTCAGATACAGAGGGAGAGAGCCAAACATGTTTACAATATAAGTGTACAGACCCCT 300
QY 301 TCTGAAAAATCCTTGCTTTAGATCTCTCTCAACACATCCGCTGACTATCCGAAATGCAAA 360
Db 301 TCTGAAAAATCCTTGCTGCTGATCCTCCAGTATATATCCGATTTACCTTAATGTAA 360
QY 361 ACTATCCATCATATTCAGAGTCAAAAACCTCATGACAGGGGATCGCCCTTCATTATTGG 420
Db 361 ACTGTTTCATCATATTCAGGTCAAAAACCTCATGACAGGGGATTCGCCCTTCATTATTGG 420
QY 421 GGAGCATTTTTCTGTATGATCGCATTTGCCCTCCACAAATGTACCGAGGCAAACTCTC 480
Db 421 GGAGCATTTTTCTGTATGATCGCGTTCCTCTACAATGTACCGAGGCAAACTCTC 480
QY 481 ACTGAAGGAAACATAGCAGCTATGATTTCTAATAAGACAGATGCAAAAATGATTTTCTCG 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATTTCTAATAAGACAGATTCACAGATGATTTTCT 540
QY 541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTTACTATAAATATTTGGACAAGT 600
Db 541 AGGCAAGGACAAAGGTATTCGTACATGAACCTTGACCTCCACCAATATAATTTGGCAAGC 600
QY 601 AGTAACGGAACGCAACGAATGACACTGGATGTTTCGGGCTCTTCAAGATACAAATCT 660
Db 601 AGCAATGAACGCAAGAAATGATACGGGATGTTTTCGATCTCTCAAGATACAAATCT 660
QY 661 ACAAGAGAACCAACATGTGCTCCGTCGTCACAAATATCTCCACCACTGCCACAGCCCGTCCG 720
Db 661 ACAACATCAACATGATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG 720
QY 721 GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780

Db	721	AGCATT	CAC	TCT	TACA	AAAT	TACT	CAAA	TAA	TACT	GTCT	AAAT	CT	TGAA	ACT	TGGA	ACT	TAT	GA	CCCA	AGT	780
Qy	781	AGTGAT	GAT	GAGG	ACCT	CGCAA	CAT	CGCG	TCT	CAGGG	TCCG	GAGAA	CGAGAA	CCCC	CACACA	840						
Db	781	AGCGAG	AT	GAGG	ACCT	TAT	GTAT	TTC	CGGG	TCT	CAGGAT	CT	TGAG	AGAC	GAGGGG	CCCC	CACACA	840				
Qy	841	ACTTCT	GAT	TGGG	TCA	CA	AGCA	AGG	CGCT	TTT	CA	TCAA	CAAT	TGCA	CCAC	CCCT	CA	CCCA	900			
Db	841	ACTCTT	AT	TGT	AGT	CA	CT	GAAC	AGAA	CAAT	CGT	CA	CAAT	TAT	TGTC	CACT	CTCT	CACTA	900			
Qy	901	CAACCA	AGCA	CGCC	ACAG	CAAGG	AGAA	CAAC	ACAA	ACCA	TTC	CCCA	AGAT	GTCT	GTG	ACT	960					
Db	901	CATCCA	AGCAC	CT	CACA	CAAT	GAGCA	AAAA	CAGT	ACGA	ATCT	CTTCC	GACA	TGCT	GTAA	ACT	960					
Qy	961	GAACTA	GACAAA	ATA	CACA	ACT	TGCA	CAAC	CGGT	CCAT	GTCC	ATG	CCCT	CTAT	ACACT	TAC	CA	CA	1020			
Db	961	GAGCA	AT	TGNA	CCG	ACCC	CNA	CAAC	ACAC	CGAC	ACG	TCT	CTCA	CAT	ACT	ATA	TACA	1020				
Qy	1021	ATCTCT	ACT	AACA	CA	CCT	CCAAA	CA	CAACT	T	CAGCA	CTCT	CTCT	GTG	CA	CAAT	TACAAA	1080				
Db	1021	ACTCCC	ACCT	TAT	AACA	CTCT	CAAGT	ACA	CACT	CAGT	ACTCT	CTTCCC	CTCC	CAAC	CCGCA	AC	1080					
Qy	1081	ACACCA	AT	TGCA	CA	CA	AGCA	CA	CAAT	CA	CTGAAAA	AT	GAGCA	AAAC	CGAT	GTCC	CTCG	1140				
Db	1081	ATCACCA	AT	AT	GATA	CACA	CA	CGT	GTAA	CT	TAGC	AGAA	CGCA	CAAA	CAAT	GTCT	CA	TTG	1140			
Qy	1141	ATAACA	CCCT	CGCT	CCAA	CGG	AAAT	TCC	CA	CA	CGAC	CAAG	AGCA	CCAG	CA	CAAAAA	1200					
Db	1141	AACA	CA	ACT	CT	GAT	TC	CA	CA	AAAT	TCC	CA	CA	CGAC	CA	CA	CA	CA	1200			
Qy	1201	GGCCCC	CCCA	CA	CGG	CA	CA	CA	CA	CA	CA	CA	CA	CA	CA	CA	CA	CA	1260			
Db	1201	AACAT	CAT	CAT	GAC	GA	CAT	CAGAT	TATA	CAAG	CAAA	CAC	CCCA	CAAA	TCTT	CTCC	GGAT	1260				
Qy	1261	CCGAGCT	CGAT	GTCA	CA	CA	AT	CT	TGT	TAT	TTT	CAGA	AGAA	AGGAT	CGAT	TTTT	CTG	AAA	1320			
Db	1261	TCTAGT	CCGA	CA	AC	CCG	CCCT	CTCT	ATA	TACT	TTAG	AA	AGAA	CGA	AGCA	TTTT	CTG	AAA	1320			
Qy	1321	GAAGGT	GAT	TAT	T	CCCG	TTTT	TAGAT	GGGT	TAA	TAA	TA	ACT	CGAA	TTG	AT	TTGAT	TCA	1380			
Db	1321	GAAGGT	GAT	TAT	T	CCCG	TTTT	TAGAT	GGGT	TAA	TAA	TA	ACT	CGAA	TTG	AT	TTGAT	TCA	1380			
Qy	1381	ATCCAA	ACAC	AGAA	CA	AT	CTTT	TGAT	GAAT	CTCC	CA	AGCT	TTA	AT	ACT	CA	CA	TAAT	TAG	1440		
Db	1381	ATCCAA	ACAC	AGAA	CA	AT	CTTT	TGAT	GAAT	CTCC	CA	AGCT	TTA	AT	ACT	CA	CA	TAAT	TAG	1440		
Qy	1441	GAACAA	CAC	ACT	CCCC	GAAT	AT	CA	GT	TAA	CTTT	CTCT	TAT	TTT	CTG	ATA	AAAA	TGGA	1500			
Db	1441	GAACAA	CAC	ACT	CCCC	GAAT	AT	CA	GT	TAA	CTTT	CTCT	TAT	TTT	CTG	ATA	AAAA	TGGA	1500			
Qy	1501	GAT	ACT	CGCT	ACT	CTG	GGG	AAAC	CGAA	TGA	TTG	TGAT	CGAC	AGT	TGAG	AT	TTCG	AGT	1560			
Db	1501	GAT	ACT	CGCT	ACT	CTG	GGG	AAAC	CGAA	TGA	TTG	TGAT	CGAC	AGT	TGAG	AT	TTCG	AGT	1560			
Qy	1561	GTG	CAG	GAG	CA	CGAT	TTT	GG	CGG	CAG	GGCT	TAG	CT	TG	GAT	CA	CA	TTTT	TG	CCCT	CGAATC	1620
Db	1561	GTG	CAG	GAG	CA	CGAT	TTT	GG	CGG	CAG	GGCT	TAG	CT	TG	GAT	CA	CA	TTTT	TG	CCCT	CGAATC	1620
Qy	1621	GAAGGA	CT	CT	AT	ACT	CG	CGGT														

Db 1801 ACATGCAAGGTGCTAGGACCTGATGTTGTCATAGGAATAGAAAGATCTATCTTAAAAATATC 1860

Qy 1861 TCAGAACAAATCGACAAAATCAGAAAGAGATGACAAACAAAAGAGAGAAACTGGCTGGGGTCTA 1920

Db 1861 TCAGAACAAATCGACAAAATCAGAAAGAGATGACAAACAAAAGAGAGAAACTGGCTGGGGTCTA 1920

Qy 1921 GGTGGCAAAATGGTGGACATCTGACTGGGGTGTCTCAACCAATTTGGGCATCTCGTACTA 1980

Db 1921 GGTGGCAAAATGGTGGACATCTGACTGGGGTGTCTCAACCAATTTGGGCATCTCGTACTA 1980

Qy 1981 TTATCTATAGCTGTTCTGATGTCCTCTCTCGTATATCTGTCGTATCTTCACTTAAATACATT 2040

Db 1981 TTATCTATAGCTGTTCTGATGTCCTCTCTCGTATATCTGTCGTATCTTCACTTAAATACATT 2040

Qy 2041 GGATGA 2046

Db 2041 GGATGA 2046

RESULT 14

AX171719

LOCUS AX171719 2046 bp DNA linear PAT 15-APR-2003

DEFINITION Sequence 7 from Patent W002079239.

ACCESSION AX171719

VERSION .

KEYWORDS AX171719.1 GI:29890728

SOURCE synthetic construct

ORGANISM synthetic construct

other sequences; artificial sequences.

1

REFERENCE

AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.

TITLE Chimeric filovirus glycoprotein

JOURNAL Patent: WO 02079239-A 7 10-OCT-2002;

U.S. Army Medical Research Institute of Infectious Diseases (US)

FEATURES

source

1. .2046

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="chimeric molecule between Marburg virus strain

Raven Glycoprotein 1 and Marburg virus strain Musoke

Glycoprotein 2"

ORIGIN

Query Match 64.5%; Score 1319.6; DB 6; Length 2046;

Best Local Similarity 77.8%; Pred No. 0;

Matches 1592; Conservative 0; Mismatches 454; Indels 0; Gaps 0;

Qy 1 ATGAAGACCAATGTTTCTTATCAGTCTTTATCTTAATTCAGGGACAAAATATCTCCCC 60

Db 1 ATGAAGACCAATATTTTCTGATTAGTCTCATTTTATCCAAAGTATAAAACTCTCCCT 60

Qy 61 ATTTTGAAGATAGCTAGTATAATCAACCCCAAAATGTGGAATCGGTATGCTCCGGAAC 120

Db 61 GTTTTGAAGATTTGCTAGTAAACAGCAACCTCAAGATGTAGATTCACTGTCGTCGGAACC 120

Qy 121 CTCAGAAAGACAGAAAGCGTCCATCTCATGCGGATTCACACTGAGTGGGCAAAAAGTTGCT 180

Db 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGGATTTACCTGAGTGGGCAAAAAGTTGCT 180

Qy 181 GATTCCCTTTGGAGGACATCCAAACGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240

Db 181 GATTCCCTTTGGAGGACATCTAAACGATGGGCTTTTCAGGACAGGTGTTCCTCCCAAGNA 240

Qy 241 GTTGAGTACACAGAGGGGGAGGAAGCCAAAACATGCTCAATATAAGTGTAAACGATCCC 300

Db 241 GTTGAGTATACGGAAGGAGAGAAAGCCAAAACATGTTTCAATATAAGTGTAAACAGACCC 300

Qy 301 TCTCGAAAATCCTTGTCTGTTAGATCCTCTTACCAACATCCGTGACTATCCGAATGCAA 360

Db 301 TCTCGAAAATCCTTGTCTGCTGGATCTCTCCCAAGTAAATCCCGGATTAACCTTAATGTA 360

Qy 361 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTATG 420

[illegible]

Db	1441	GATCAACATGCGCTCCCCCAATATTAGTTAACTTTTAACTTTTCTTCTAATAATAATGAG	1501			
Qy	1501	GATACTGCCCTACTCTCGGGGAAAACGAGAATGATTGTGATGCGAGATTTGAGGATTTGGAGT	1560			
Db	1501	AAACATGGCTACTCTCGAGAAATGAGATGATTGTGATGCGAGATTTAAGAAATTTGGAGC	1560			
Qy	1561	GTCAGAGGAGCAATTTGGCGGCAGGGCTTAGCTGGAATACCAATTTTGTGGCCCTCGAATC	1620			
Db	1561	GTTTCAGGAGGATGACCTGGCCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTCGAAAT	1620			
Qy	1621	GAGGGACTCTATACTCGCGGTTTAAATCAAAATCAAGACAATTTAGTTGTAGGTTTGAGG	1680			
Db	1621	GAAGGACTTTTACACTGCTGTTTTTAAATTAATAATCAAAACAATTTGGTCTGCAGGTTGAGG	1680			
Qy	1681	CGCTTAGCTAATCAAACTGCTAAATCCTTGGAGCTCTTGTTAAGGGTCACAACCGAGGAA	1740			
Db	1681	CGTCTAGCCCAATCAAACTGCCAATCCTTTGGAATCTTATTGAGAGTCACAACTGAGGAA	1740			
Qy	1741	AGGACATTTTCCTTAATCAATAGGCATGCAATTTGACTTTTTTGTCTTACGAGGTGGGCGGA	1800			
Db	1741	AGAACAATTCCTTAATCAATAGACATGCTATTGACTTTTCTACTCAAGATGGGAGGA	1800			
Qy	1801	ACATGCNAGTGTGAGCACTGATTTGTCATAGGAATAGAAGATCTATCTAAAAATATC	1860			
Db	1801	ACATGCAAGTGTGTGGACCTGATTTGTTCATCGGATAGAAGACTTGTCCAAAAATATT	1860			
Qy	1861	TCAGAACAAATCGACAAAATCAGAAAGGATGAACAAAAGGAGGAAAACCTGGCTGGGGTCTA	1920			
Db	1861	TCAGAGCAATTTGACCAATTTAAAGAGGACGAACAAAAGAGGGGACTTGTGTGGGTTCTG	1920			
Qy	1921	GGTGGCAAAATGTTGGACATCTGACTGGGTGTTTCTCACCAATTTGGGCATCTCTGCTACTA	1980			
Db	1921	GGTGGTAAATGGTGGACATCCGACTGGGTGTTTCTTACTTAATCTGGGCAATTTTGTCTACTA	1980			
Qy	1981	TTATCTATAGCTGTTCTGATGCTCTGCTCTCTGATCTGTCGTATCTTCATCAATAATACATT	2040			
Db	1981	TTATCCATAGCTGTTCTGATTGCTCTATCCTCTGATTTGTGCTATCTTTACTAAATATATC	2040			
Qy	2041	GGATGA 2046				
Db	2041	GGATAA 2046				
RESULT 15						
AX717715						
LOCUS	AX717715	1841 bp	DNA linear PAT 15-APR-2003			
DEFINITION	Sequence 3 from Patent WO02079239.					
ACCESSION	AX717715					
VERSION	AX717715.1 GI:29890726					
KEYWORDS	synthetic construct					
SOURCE	synthetic construct					
ORGANISM	other sequences; artificial sequences.					
REFERENCE	1					
AUTHORS	Grogan,C.C., Hevey,M.C. and Schmaljohn,A.L.					
TITLE	Chimeric filovirus glycoprotein					
JOURNAL	Patent: WO 02079239-A 3 10-OCT-2002;					
FEATURES	U.S. Army Medical Research Institute of Infectious Diseases (US)					
	Location/Qualifiers					
	1. .1841					
	/organism="synthetic construct"					
ORIGIN	/mol_type="unassigned DNA"					
	/db_xref="taxon:32630"					
	/note="chimeric molecule between Marburg virus strain					
	Mubeke Glycoprotein 1 and Ebola virus Zaire Mayinga strain Glycoprotein 2"					
Query Match 63.9%; Score 1308; DB 6; Length 1841;						
Best Local Similarity 100.0%; Pred. No. 0;						
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Ov	1	ATGAAGACCACATGTTTCCCTTATCAGTCTTATCTTATTTCAAGGAGGACAAAATAATCTCCCC	60			

ORIGIN

Query Match 63.9%; Score 1308; DB 6; Length 1841;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```
Db      |||||
1. ATGAAGACACATGTTTCCTTATCAGTCTTATCTTTAATTCAGGGCAAAAAATCTCCCC 60
QY      |||||
61 ATTTTAGAGATAGCTAGTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db      |||||
61 ATTTTAGAGATAGCTAGTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
QY      |||||
121 CTCCAGAGACAGAGACGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAGTGTCT 180
Db      |||||
121 CTCCAGAGACAGAGACGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAGTGTCT 180
QY      |||||
181 GATTCCTCTTTGGAGGCATCCAAGGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
Db      |||||
181 GATTCCTCTTTGGAGGCATCCAAGGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
QY      |||||
241 GTTGTAGTACAGAGGGGAGGAAGCCAAAACATGCTACATATATAGTGTAAAGGATCCC 300
Db      |||||
241 GTTGTAGTACAGAGGGGAGGAAGCCAAAACATGCTACATATATAGTGTAAAGGATCCC 300
QY      |||||
301 TCTGGAAAATCCTTCTGTAGATCCTCTCTACCAACATCCGTGACTATCCGNAATGCAAA 360
Db      |||||
301 TCTGGAAAATCCTTCTGTAGATCCTCTCTACCAACATCCGTGACTATCCGNAATGCAAA 360
QY      |||||
361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACACAGGGATCGCCCTTCATTTATGG 420
Db      |||||
361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACACAGGGATCGCCCTTCATTTATGG 420
QY      |||||
421 GGAGCATTTTTCTGTATGATGCGATGCTCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
Db      |||||
421 GGAGCATTTTTCTGTATGATGCGATGCTCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
QY      |||||
481 ACTGAAGGGAACATAGCAGCTATGATGCTCAATAGACAGTGCACAAAATGATTTCTCG 540
Db      |||||
481 ACTGAAGGGAACATAGCAGCTATGATGCTCAATAGACAGTGCACAAAATGATTTCTCG 540
QY      |||||
541 CGGCAAGGACAGGGTACCGTCAATGAATCTGACTTCTACTATAAATATTGGACAAGT 600
Db      |||||
541 CGGCAAGGACAGGGTACCGTCAATGAATCTGACTTCTACTATAAATATTGGACAAGT 600
QY      |||||
601 AGTAAACGGAACGCAAAACGAATGACACTGGATGTTTCGGGCTCTTCAAGAATACAAATCT 660
Db      |||||
601 AGTAAACGGAACGCAAAACGAATGACACTGGATGTTTCGGGCTCTTCAAGAATACAAATCT 660
QY      |||||
661 ACAAGAACCAAAACATGTGCTCCGTCCAAATACCTCCACACATGCGCCACAGCCGTCGG 720
Db      |||||
661 ACAAGAACCAAAACATGTGCTCCGTCCAAATACCTCCACACATGCGCCACAGCCGTCGG 720
QY      |||||
721 GAGATCAAACTCACAGACACCCCACTGATGCCACCAAACTCAATACCACGAGCCCAAGC 780
Db      |||||
721 GAGATCAAACTCACAGACACCCCACTGATGCCACCAAACTCAATACCACGAGCCCAAGC 780
QY      |||||
781 AGTGTATGATGAGGACTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
Db      |||||
781 AGTGTATGATGAGGACTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
QY      |||||
841 ACTTCTGTATGCGGTACCAAGGAGGGCTTTCATCAATGCGCACCCACTCCCTCACCA 900
Db      |||||
841 ACTTCTGTATGCGGTACCAAGGAGGGCTTTCATCAATGCGCACCCACTCCCTCACCA 900
QY      |||||
901 CAACCAAGCAGCCACAGCAGGAGGAAACACACAAACCATTTCCCAAGATGCTGTGACT 960
Db      |||||
901 CAACCAAGCAGCCACAGCAGGAGGAAACACACAAACCATTTCCCAAGATGCTGTGACT 960
QY      |||||
961 GAACTAGACAAAATAACAACTCCACAAACCGTCCATGCCCTCTCATACACTACCACA 1020
Db      |||||
961 GAACTAGACAAAATAACAACTCCACAAACCGTCCATGCCCTCTCATACACTACCACA 1020
QY      |||||
1021 ATCTCTATCAACAACTCCCAACCAAACTTCAGCACTCTCTGTGCAACCAATTAACAAAC 1080
Db      |||||
1021 ATCTCTATCAACAACTCCCAACCAAACTTCAGCACTCTCTGTGCAACCAATTAACAAAC 1080
QY      |||||
1081 ACCACCAATGACACACAGAGCAATCACTGAAAATGAGCAACACAGTGCCTCCCTCG 1140
Db      |||||
```

```
Db      |||||
1081 ACCACCAATGACACACAGAGCAATCACTGAAAATGAGCAACACAGTGCCTCCCTCG 1140
QY      |||||
1141 ATAAACAACCTCGCTCCAAACGGGAAATCCCAACCAAGAGCAACAGAGCAAAAAA 1200
Db      |||||
1141 ATAAACAACCTCGCTCCAAACGGGAAATCCCAACCAAGAGCAACAGAGCAAAAAA 1200
QY      |||||
1201 GGGCCCGCCACAACGGCACCAACACAGCAAAATGAGCATTTTACCAGTCTCTCCCTCC 1260
Db      |||||
1201 GGGCCCGCCACAACGGCACCAACACAGCAAAATGAGCATTTTACCAGTCTCTCCCTCC 1260
QY      |||||
1261 CCCAGCTCGACTGCACACATCTTGTATATTTTCAGAGAAAGCGATCG 1308
Db      |||||
1261 CCCAGCTCGACTGCACACATCTTGTATATTTTCAGAGAAAGCGATCG 1308
```

Search completed: August 1, 2005, 02:43:48
Job time : 8800.18 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:55:23 ; Search time 1071.87 Seconds
(without alignments)
11299.693 Million cell updates/sec

Title: US-10-066-506A-5

Perfect score: 2046

Sequence: 1 atgaagaccacatgttcct.....tcactaaatacatggatga 2046

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2002bs.*

8: geneseqn2003as.*

9: geneseqn2003bs.*

10: geneseqn2003cs.*

11: geneseqn2003ds.*

12: geneseqn2004as.*

13: geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2046	100.0	2046	10	ABT13452 DNA encod
2	1841.2	90.0	2103	8	ACC57764 Marburg v
3	1841.2	90.0	2164	4	AAD04042 Marburg v
4	1841.2	90.0	7778	10	ACC71550 VRC6701 (
5	1841.2	90.0	11460	3	Acc71550 VRC6701 (
6	1825.2	89.2	2100	3	Az87211 VEE repli
7	1813.6	88.6	2051	10	Az51039 Marburg v
8	1767.6	86.4	7005	10	ABT13454 DNA encod
9	1767.6	86.4	8256	10	ACC71551 VRC6702 (
10	1692.4	82.7	19112	13	Acc71552 VRC6710 (
11	1518	74.2	2046	10	Adm48345 Marburg v
12	1510	73.8	2247	4	ABT13456 DNA encod
13	1313.6	64.5	2046	10	Ad04043 Marburg v
14	1308	63.9	1841	10	ABT13453 DNA encod
15	995.6	48.7	6902	10	ABT13451 DNA encod
16	546.6	26.7	2252	10	ACC71552 VRC6703 (p
17	145	7.1	6940	10	ABT13450 DNA encod
18	145	7.1	7073	2	Acc71535 VRC6202 (
19	145	7.1	7082	10	Aax59391 Plasmid p
20	145	7.1	7087	10	ACC71533 VRC6200 (
					Acc71534 VRC6201 (

21	145	7.1	8221	10	ACC71547
22	145	7.1	8338	10	ACC71546
23	145	7.1	10783	10	ACC71545
24	143.4	7.0	7073	2	AAX89796
25	140.4	6.9	7001	2	AAX89795
26	138.8	6.8	6885	10	ACC71538
27	138.8	6.8	6889	10	ACC71539
28	138.8	6.8	7002	10	ACC71536
29	138.8	6.8	7003	2	AAX59390
30	138.8	6.8	7023	10	ACC71541
31	138.8	6.8	7036	10	ACC71537
32	138.8	6.8	8146	10	ACC71540
33	131.2	6.4	2039	10	ABT13455
34	131.2	6.4	2172	4	AAD04041
35	131.2	6.4	2224	3	AAX51038
36	131.2	6.4	2298	3	AAX87189
37	131.2	6.4	2298	4	AAF76953
38	131.2	6.4	2298	12	ADL27460
39	131.2	6.4	6467	10	ACC71530
40	131.2	6.4	6561	10	ACC71524
41	131.2	6.4	6624	10	ACC71523
42	131.2	6.4	6914	10	ACC71529
43	131.2	6.4	7154	10	ACC71521
44	131.2	6.4	7188	10	ACC71522
45	131.2	6.4	7272	2	AAX59393

ALIGNMENTS

RESULT 1

ABT13452
ID ABT13452 standard; DNA; 2046 BP.

XX ABT13452;

XX 30-JAN-2003 (first entry)

XX DNA encoding a chimeric filovirus protein MUS-GP1/RVN-GP2.
XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
XX immune response; Ebola; Marburg virus; vaccine; gene; ds.

XX Unidentified.

XX Chimeric.

XX WO200279239-A2.

XX 10-OCT-2002.

XX 31-JAN-2002; 2002WO-US003339.

XX 31-JAN-2001; 2001US-0267522P.

XX (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.

XX Grogan CC, Hevey MC, Schmaljohn AL;

XX WPI; 2003-040651/03.

XX P-FSDB; ABJ18474.

XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
XX useful for inducing an immune response against infection of different
XX filoviruses, specifically against both Ebola and Marburg viruses.
XX Claim 16; Page 74-76; 94pp; English.
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
XX comprising GP1 and GP2, where GP1 is from a filovirus different than that
XX of GP2. The chimeric filovirus GP protein is useful for inducing an
XX immune response against infection of different filoviruses, specifically
XX against both Ebola and Marburg viruses by being used as a vaccine. This
XX polynucleotide sequence represents a DNA encoding a chimeric filovirus

CC	protein of the invention	
XX		
SQ	Sequence 2046 BP; 659 A; 517 C; 402 G; 468 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 2046; DB 10; Length 2046;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 2046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGAAGACCAATGTTTCCCTTATCAGTCTTATCTTAAATCAAGGGACAAAAAATCTCCCC	60
DB	1 ATGAAGACCAATGTTTCCCTTATCAGTCTTATCTTAAATCAAGGGACAAAAAATCTCCCC	60
QY	61 ATTTTAGAGTAGTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	120
DB	61 ATTTTAGAGTAGTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	120
QY	121 CTCAGAGACAGAGACGTCTCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT	180
DB	121 CTCAGAGACAGAGACGTCTCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT	180
QY	181 GATTTCCCTTTTGGAGCATCCAAAGCATGGGCTTTTCAAGCAGGTGTACTCCCAAGAAAT	240
DB	181 GATTTCCCTTTTGGAGCATCCAAAGCATGGGCTTTTCAAGCAGGTGTACTCCCAAGAAAT	240
QY	241 GTTGAGTACACAGAGGGGAGGAAGCCAAACATGCTACAATATAAGTGTAAACGGATCCC	300
DB	241 GTTGAGTACACAGAGGGGAGGAAGCCAAACATGCTACAATATAAGTGTAAACGGATCCC	300
QY	301 TCTGGAAATCCTTGCTGTAGATGCTCTTACCAACATCCGTGACATATCCGAAATGCAAA	360
DB	301 TCTGGAAATCCTTGCTGTAGATGCTCTTACCAACATCCGTGACATATCCGAAATGCAAA	360
QY	361 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	420
DB	361 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	420
QY	421 GGAGCAATTTTTCTGTATGATCGCATGCTCCCAACAAATGTACCGAGGCAAAAGTCTTC	480
DB	421 GGAGCAATTTTTCTGTATGATCGCATGCTCCCAACAAATGTACCGAGGCAAAAGTCTTC	480
QY	481 ACTGAAGGGAACATAGCAGCTATGATGTCAATATAGACAGTGCACAAATGATTTTCTCG	540
DB	481 ACTGAAGGGAACATAGCAGCTATGATGTCAATATAGACAGTGCACAAATGATTTTCTCG	540
QY	541 CGGCAAGGACAAAGGTACCGTCAATGAAATCTGACTTCTACTAAATAATTTGGACAAGT	600
DB	541 CGGCAAGGACAAAGGTACCGTCAATGAAATCTGACTTCTACTAAATAATTTGGACAAGT	600
QY	601 AGTAACGGGAACGCAACGAAATGACACTGGATGTTTCGGCGCTCTTCAAGAATACAAATCT	660
DB	601 AGTAACGGGAACGCAACGAAATGACACTGGATGTTTCGGCGCTCTTCAAGAATACAAATCT	660
QY	661 ACAAGAACCAACATGCTCGTCCGTCACAAATACCTCAGCACTGCCACAGCCCGTCCG	720
DB	661 ACAAGAACCAACATGCTCGTCCGTCACAAATACCTCAGCACTGCCACAGCCCGTCCG	720
QY	721 GAGATCAAACTCACAGACCCCAACTGATGCCACCAACTCAATACCAAGGACCAAGC	780
DB	721 GAGATCAAACTCACAGACCCCAACTGATGCCACCAACTCAATACCAAGGACCAAGC	780
QY	781 AGTGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACCGGCAACACA	840
DB	781 AGTGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACCGGCAACACA	840
QY	841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCAATCAAAATGCGCACTCCCTCAGCA	900
DB	841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCAATCAAAATGCGCACTCCCTCAGCA	900
QY	901 CAACCAAGCACGCCACAGCAAGGAGAACACAAACCAATTCCTCAAGATGCTGTGACT	960
DB	901 CAACCAAGCACGCCACAGCAAGGAGAACACAAACCAATTCCTCAAGATGCTGTGACT	960
QY	961 GAACTAGACAAAAATAACA CAACTGCAACCGTCCATGCCCCCTCATAACTACCCACA	1020
961	GAACTAGACAAAAATAACA CAACAACTGCAACCGTCCATGCCCCCTCATAACTACCCACA	1020
1021	ATCTCTACTTAAACACACCTCCAAACACAACTTCAAGCACTCTCTGACCACTTACAAAC	1080
1021	ATCTCTACTTAAACACACCTCCAAACACAACTTCAAGCACTCTCTGACCACTTACAAAC	1080
1081	ACCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1140
1081	ACCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1140
1141	ATACCAACCTTCCCTCCAAACCGGAAATCCCAACAGCAAGCAAGCAAGCAAGCAAGCA	1200
1141	ATACCAACCTTCCCTCCAAACCGGAAATCCCAACAGCAAGCAAGCAAGCAAGCAAGCA	1200
1201	GGCCCGGCAACAAAGGCAACCAACAGCAAAATGAGCAATTTTCAACAGTCTCTCCCCACC	1260
1201	GGCCCGGCAACAAAGGCAACCAACAGCAAAATGAGCAATTTTCAACAGTCTCTCCCCACC	1260
1261	CCAGCTCGACTGCAACAACTCTTTGTATATTTTCAAGAAAGCGATCGATTTCTGGAAA	1320
1261	CCAGCTCGACTGCAACAACTCTTTGTATATTTTCAAGAAAGCGATCGATTTCTGGAAA	1320
1321	GAAAGTGATATATTTCCCGTTTATAGATGGGTTAATAAATCTGAAATTTGATTTGATCCA	1380
1321	GAAAGTGATATATTTCCCGTTTATAGATGGGTTAATAAATCTGAAATTTGATTTGATCCA	1380
1381	ATCCCAACACAGAAACAACTCTTTGATGAATCTCCAGCTTTTAACTTCACTTAATGAG	1440
1381	ATCCCAACACAGAAACAACTCTTTGATGAATCTCCAGCTTTTAACTTCACTTAATGAG	1440
1441	GAAACACACACTCCCGGAAATCAAGTTTAACTTTCTCTTATTTTCTGATAAAATGGA	1500
1441	GAAACACACACTCCCGGAAATCAAGTTTAACTTTCTCTTATTTTCTGATAAAATGGA	1500
1501	GATCTGCTACTCTGCGGAAAAAAGAGATGATTTGATGACAGAGTTGAGGATTTGGAGT	1560
1501	GATCTGCTACTCTGCGGAAAAAAGAGATGATTTGATGACAGAGTTGAGGATTTGGAGT	1560
1561	GTGAGGAGGAGCAATTTGCGCGAGGCTTACGTGATACCAATTTTGGCCCTCGAATC	1620
1561	GTGAGGAGGAGCAATTTGCGCGAGGCTTACGTGATACCAATTTTGGCCCTCGAATC	1620
1621	GAAAGCACTATCTACTGCGGTTTAAATCAAAATCAGAAACAATTTAGTTTGTAGGTTGAG	1680
1621	GAAAGCACTATCTACTGCGGTTTAAATCAAAATCAGAAACAATTTAGTTTGTAGGTTGAG	1680
1681	CGTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGTCACAAACGAGGAA	1740
1681	CGTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGTCACAAACGAGGAA	1740
1741	AGGACATTTTCTTAACTAATAGCATGCAATTTGATTTTTCCTTACGAGGTGGGCGGA	1800
1741	AGGACATTTTCTTAACTAATAGCATGCAATTTGATTTTTCCTTACGAGGTGGGCGGA	1800
1801	ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAATCTATCTAATAATATC	1860
1801	ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAATCTATCTAATAATATC	1860
1861	TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGTCTA	1920
1861	TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGTCTA	1920
1921	GGTGGCAATGGTGACATCTGATGGGGTCTTCTCACCATAATTTGGGATCTCGTACTA	1980
1921	GGTGGCAATGGTGACATCTGATGGGGTCTTCTCACCATAATTTGGGATCTCGTACTA	1980
1981	TTATCTATAGCTGTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2040
1981	TTATCTATAGCTGTTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2040
2041	GGATGA 2046	

Db 2041 GGATCA 2046

RESULT 2
ACC57764
ID ACC57764 standard; cDNA; 2103 BP.
XX
AC ACC57764;
XX
DT 28-JUL-2003 (first entry)
XX
DE Marburg virus envelope glycoprotein coding sequence.
XX
KW Lentivirus; pseudotyped virus; gene therapy; vector; gene; ss.
XX
OS Marburg virus.
XX
FH Key Location/Qualifiers
FT CDS 10..2055
FT /*tag= a
FT /product= "Envelope glycoprotein"
XX
XX WO2003035849-A2.
XX
PD 01-MAY-2003.
XX
XX 28-OCT-2002; 2002WO-US034545.
XX
XX 26-OCT-2001; 2001US-0353221P.
XX
XX 26-OCT-2001; 2001US-0356436P.
XX
XX (MCCR/) MCCRAY P B.
XX
XX (SAND/) SANDERS D A.
XX
XX (DAVI/) DAVIDSON B L.
XX
XX McCreay PB, Sanders DA, Davidson BL;
XX
XX WPI; 2003-421416/39.
XX
XX P-PSDB; ABR42242.
XX
XX New pseudotyped lentivirus comprising a lentiviral capsid, a lipid bilayer and a Marburg glycoprotein disposed in the lipid bilayer, useful for eliciting an immune response against feline immunodeficiency virus infection.
XX
XX Disclosure; Page 36-39; 41pp; English.
XX
XX The present sequence is the coding sequence for the envelope glycoprotein of Marburg virus. The invention provides methods for gene transfer to cells using glycoprotein-pseudotyped lentiviruses. The glycoprotein is preferably Marburg glycoprotein. The Marburg viral glycoprotein preferably has a mutation in the C-terminal portion (see ABR42243-47) that results in a higher titre production of the pseudotyped virus. The pseudotyped lentivirus is used to introduce nucleic acid sequences encoding a desired protein into a hepatocyte, brain glial or airway epithelial cell, in vitro or in vivo. The desired protein is preferably cystic fibrosis transmembrane conductance regulator, or is the low density lipoprotein receptor, alpha1-antitrypsin, ornithine transcarbamylase, Factor VIII or a high affinity glutamate receptor (all claimed)
XX
XX Sequence 2103 BP; 671 A; 533 C; 408 G; 491 T; 0 U; 0 Other;
XX
XX Query Match 90.0%; Score 1841.2; DB 8; Length 2103;
XX
XX Best Local Similarity 93.7%; Pred No. 0;
XX
XX Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
XX
XX 1 ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
XX
XX 10 ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 69
XX
XX 61 ATTTAGAGATAGCTAGTAAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACCT 120
XX
XX

Db 70 ATTTAGAGATAGCTAGTAAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACCT 129
Qy 121 CTCAGAAAGACAGAAAGAGCTCCATCTCATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
Db 130 CTCAGAAAGACAGAAAGAGCTCCATCTCATGGGATTCACATGAGTGGGCAAAAAGTTGCT 189
Qy 181 GATTCCCTTTGGAGGATCCAGCGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 240
Db 190 GATTCCCTTTGGAGGATCCAGCGGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 249
Qy 241 GTTCAGTACACAGAGGGGAGGAAAGCAAAACATGTCAATATAAGTGTAAAGGATCCCC 300
Db 250 GTTCAGTACACAGAGGGGAGGAAAGCAAAACATGTCAATATAAGTGTAAAGGATCCCC 309
Qy 301 TCTGAAAAATCCTTTGCTTTAGATCCTCTCAACCAATCCCTGATATCCGAAATGCAAAA 360
Db 310 TCTGAAAAATCCTTTGCTTTAGATCCTCTCAACCAATCCCTGATATCCGAAATGCAAAA 369
Qy 361 ACTATCCATCATATTCAGGTCAAAACCCCTCATGTCAGAGGGATCGCCCTTCATTTATGG 420
Db 370 ACTATCCATCATATTCAGGTCAAAACCCCTCATGTCAGAGGGATCGCCCTTCATTTATGG 429
Qy 421 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 480
Db 430 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACAAATGTACCGAGGCAAAAGTCTTC 489
Qy 481 ACTGAAGGGAAACATAGCAGCTATGATTTGTCAATGAAGACAGTGCACAAATGATTTTCG 540
Db 490 ACTGAAGGGAAACATAGCAGCTATGATTTGTCAATGAAGACAGTGCACAAATGATTTTCG 549
Qy 541 CGGCAAGGACAAAGGTACCGTTCATATGANTCTGACTTCTACTAATTAATATGGACAGT 600
Db 550 CGGCAAGGACAAAGGTACCGTTCATATGAATCTGACTTCTACTAATTAATATGGACAGT 609
Qy 601 AGTAAACGGAACGCAAAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAATCT 660
Db 610 AGTAAACGGAACGCAAAACGAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAATCT 669
Qy 661 ACAAGAAACCAAAACATGTGCTCCGTCCTCAAAATATCTCCACACTGCCCCACAGCCGCTCG 720
Db 670 ACAAGAAACCAAAACATGTGCTCCGTCCTCAAAATATCTCCACACTGCCCCACAGCCGCTCG 729
Qy 721 GAGATCAAACTCACAGCACCCCAACTGATGTCGACCAAACTCAATACACGAGCCCAAGC 780
Db 730 GAGATCAAACTCACAGCACCCCAACTGATGTCGACCAAACTCAATACACGAGCCCAAGC 789
Qy 781 AGTCATGATGAGGACCTCGCAACATCCGGCTCAGGTCGAGAACGAGAACCCCAACACA 840
Db 790 AGTCATGATGAGGACCTCGCAACATCCGGCTCAGGTCGAGAACGAGAACCCCAACACA 849
Qy 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTTCACCA 900
Db 850 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTTCACCA 909
Qy 901 CAACCAAGCAGCCACAGCAGAGGAGGAAACAAACCAACATTTCCCAAGATGCTGTGACT 960
Db 910 CAACCAAGCAGCCACAGCAGAGGAGGAAACAAACCAACATTTCCCAAGATGCTGTGACT 969
Qy 961 GAACCTAGCAAAAAATAACACAACCTGCAACCGTCCCTCCCTCATATAACACTACCACA 1020
Db 970 GAACCTAGCAAAAAATAACACAACCTGCAACCGTCCCTCCCTCATATAACACTACCACA 1029
Qy 1021 ATCTCTACTAAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGCACTTACAAAC 1080
Db 1030 ATCTCTACTAAACACACCTCCAAACACAACTTTCAGCACTCTCTCTGCACTTACAAAC 1089
Qy 1081 ACCACCAATGACACACAGAGACCAATCACTGAAATAGGCAAAACAGTGCCTCCCTCG 1140
Db 1090 ACCACCAATGACACACAGAGACCAATCACTGAAATAGGCAAAACAGTGCCTCCCTCG 1149
Qy 1141 ATACACACCTTCCTCCAAACGGGAAATCCCAACAGCAGCAAGAGCAGCAGGCAAAAAA 1200
Db 1150 ATACACACCTTCCTCCAAACGGGAAATCCCAACAGCAGCAAGAGCAGCAGGCAAAAAA 1209

QY 1201 GGGCCGGCCAAACGGACCAACAAACGACAAATAGCATTTCACAGTCTCTCCCCCACC 1260
DB 1210 GGGCCGGCCAAACGGACCAACAAACGACAAATAGCATTTCACAGTCTCTCCCCCACC 1269
QY 1261 CCCAGCTGCAGTGCACACATCTCTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAAA 1320
DB 1270 CCCAGCTGCAGTGCACACATCTCTGTATATTTTCAGAGAAAGCGAGTATCTCTGGAGG 1329
QY 1321 GAAGGTGATATATTCCTCGGTTTTAGATGGTTAAATAAATCTAGAAATTTGATTTGATCCA 1380
DB 1330 GAAGGGACATGTTCCCTTTCTCGATGGTTAAATAAATCTCGCAATTTGATTTGACCCA 1389
QY 1381 ATCCCAACACAGAGAAACAACTTTTGTATGAATCTCCAGCTTTTAAATCTTCACTAATGAG 1440
DB 1390 GTTCCAAATACAAACAACTTTTGTATGAATCTCTAGTTCTGTGCTCGCTCAGGAA 1449
QY 1441 GAACAACACATCCCGGAAATCTAGTTTAACTTTCTTATTTTCTGATAAATAATGA 1500
DB 1450 GATCAACATGCTCCCGCAATATAGTTTAACTTTTATCTTATTTTCTTAATATAATGAG 1509
QY 1501 GATCTGCTACTCTGGGAAACAGAGATGTTGTATGACAGATTTGAGGATTTGGAGT 1560
DB 1510 AACCTGCTACTCTGGAGAAATGAGATGATTTGTATGACAGATTTAAGAAATTTGGAGC 1569
QY 1561 GTGAGGAGGAGATTTGGCGCAGGCTTAGCTGGATACCAATTTTGGCCCTGGAATC 1620
DB 1570 GTTCAGGAGGATGACTGGCGCAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAATT 1629
QY 1621 GAAGGACTCTATCTGCGGTTTAAATCAAAAATCAGAACAAATTTAGTTGTAGTTGAGG 1680
DB 1630 GAAGGACTTTTACACTGCTGTTTTTAAATTAATAATCAAAACAAATTTGCTGCAAGTTGAGG 1689
QY 1681 CGCTTAGCTTAATCAAACTGCTAAATCCTTGGAGCTCTGTTAAGGTCACACACGAGGAA 1740
DB 1690 CGTCTAGCCAAATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 1749
QY 1741 AGGACATTTTCTTAATCAATAGGATGCAATTTGACTTTTCTTACGAGGTGGGGCGGA 1800
DB 1750 AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCAAGATGGGGAGGA 1809
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAATAATTC 1860
DB 1810 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAGACTTTGTCCAAAATATT 1869
QY 1861 TCAGAACAAATCGAACAAATCAGAAAGGATGAACAAAAGGAGGAACTGGCTGGGTCTA 1920
DB 1870 TCAGAGCAAAATGACCAAAATTAAGAGGACGACAAAAGAGGGACTGTTTGGGGTCTG 1929
QY 1921 GGTGGCAATGGTGACATCTGACTGGGTGTTCTCACCAGTTTGGGCATCTGCTACTA 1980
DB 1930 GGTGGTAATGGTGACATCTGACTGGGTGTTCTTACTAACTTGGGCATTTTGTACTA 1989
QY 1981 TTATCTATAGCTGTTCTGATGCTCTGCTGTATCTGCTGATCTTCTCACTAAATACATT 2040
DB 1990 TTATCTATAGCTGTTCTGATGCTCTGCTGTATCTGCTGATCTTTTACTAAATATATC 2049
QY 2041 GGATGA 2046
DB 2050 GGATAA 2055

RESULT 3
AAD04042
ID AAD04042 standard; DNA; 2164 BP.
XX
AC AAD04042;
XX
DT 02-JUL-2001 (first entry)
XX
DE Marburg virus Musoke strain glycoprotein (GP) DNA.

XX Glycoprotein; GP; immune response; vaccine; antiviral;

KW type I transmembrane protein; ds.
XX Marburg virus.
FH Key Location/Qualifiers
FT CDS 119..2164
FT /*tag= a
FT /product= "Marburg virus Musoke strain glycoprotein (GP)"
XX US6200959-B1.
XX 13-MAR-2001.
XX 04-DEC-1996; 96US-00760615.
XX 04-DEC-1996; 96US-00760615.
XX (POWD-) POWDERJECT VACCINES INC.
XX Haynes JR, Schmaljohn CS, Fuller DL, Schmaljohn A, Jahrling PB;
XX WPI; 2001-280564/29.
XX P-PSDB; AAE00707.
XX Inducing an immune response to a Marburg or Ebola virus involves
XX delivering a genetic vaccine, which contains a genetic construct encoding
XX antigenic determinants for filovirus, using a particle acceleration
XX device.
XX Claim 5; Col 25-30; 33pp; English.
XX The patent discloses a method of inducing an immune response to Marburg
XX or Ebola virus glycoprotein which involves delivering a genetic vaccine,
XX containing a genetic construct encoding antigenic determinants for
XX filovirus, using a particle acceleration device. The genetic vaccine for
XX filovirus is created by joining a DNA sequence encoding at least a
XX portion of the filovirus glycoprotein to a promoter effective to promote
XX transcription of the DNA sequence. This method is useful for inducing
XX humoral, cell-mediated and secretory immune responses in the treated
XX individual. The present sequence is a DNA encoding Marburg virus Musoke
XX strain glycoprotein (GP), which is a type I transmembrane protein. This
XX sequence is used in the construction of genetic vaccine against
XX filoviruses

Sequence 2164 BP; 701 A; 536 C; 412 G; 515 T; 0 U; 0 Other;
Query Match 90.0%; Score 1841.2; DB 4; Length 2164;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCC 60
DB 119 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCC 178
QY 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGGAATTCGGTATGCTCCGGAAT 120
DB 179 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGGAATTCGGTATGCTCCGGAAT 238
QY 121 CTCAGAGACAGAGAGCTCCATCTGATGGATTTCACACTGAGTGGGCAAAAATGCTGT 180
DB 239 CTCAGAGAGACAGAGAGCTCCATCTGATGGATTTCACACTGAGTGGGCAAAAATGCTGT 298
QY 181 GATTCCCCCTTTGGAGGCATCCAAGCGATGGGCTTTCCAGACAGGTGTACCTCCCAAGAT 240
DB 299 GATTCCCCCTTTGGAGGCATCCAAGCGATGGGCTTTCCAGACAGGTGTACCTCCCAAGAT 358
QY 241 GTTGAAGTACACAGAGGGGAGGAGCCAAAACATGCTCAATATATAGTTAAAGGATCC 300
DB 359 GTTGAAGTACACAGAGGGGAGGAGCCAAAACATGCTCAATATATAGTTAAAGGATCC 418
QY 301 TCTGGAAAATCCTTGTCTGTAGTCCCTTACCACATCCCGTACTATCCGAAATGCAAA 360
DB 419 TCTGGAAAATCCTTGTCTGTAGTCCCTTACCACATCCCGTACTATCCGAAATGCAAA 478

Qy	361	ACTATCCATCATATTCAAGGTCAAAAACCTCATGCA	CAGGGGATCGCCCTTCATTTATGG	420
Db	479	ACTATCCATCATATTCAAGGTCAAAAACCTCATGCA	CAGGGGATCGCCCTTCATTTATGG	538
Qy	421	GGAGCATTTTTCGTGATATGCGCATTTGCCCTCC	ACAACAATGTATACCGAGCAAGTCTTC	480
Db	539	GGAGCATTTTTCGTGATATGCGCATTTGCCCTCC	ACAACAATGTATACCGAGCAAGTCTTC	598
Qy	481	ACTGAAGGGCAACATAGCAGCTATGATTGTCAATA	GAAGACAGTGCACAAAAATGATTTCTCG	540
Db	599	ACTGAAGGGCAACATAGCAGCTATGATTGTCAATA	GAAGACAGTGCACAAAAATGATTTCTCG	658
Qy	541	CGGCAAGGACAAAGGGTACCGTCAATATGAATCT	GTACTTCTACTAATAATATTGGACAAGT	600
Db	659	CGGCAAGGACAAAGGGTACCGTCAATATGAATCT	GTACTTCTACTAATAATATTGGACAAGT	718
Qy	601	AGTAACCGGAACGCAAAACGAATGACATCGGAT	GTTTTCGGCGCTCTTCAAGGAATACAAATTC	660
Db	719	AGTAACCGGAACGCAAAACGAATGACATCGGAT	GTTTTCGGCGCTCTTCAAGGAATACAAATTC	778
Qy	661	ACAAAGAAACAAACATGTGTCCGTCCAAATACT	CCACACATGCCACACAGCCCGTCCG	720
Db	779	ACAAAGAAACAAACATGTGTCCGTCCAAATACT	CCACACATGCCACACAGCCCGTCCG	838
Qy	721	GAGATCAAACTCAACAGCACCCCAATGTATGTC	CCACCAAACTCAATACCAACCGCAACG	780
Db	839	GAGATCAAACTCAACAGCACCCCAATGTATGTC	CCACCAAACTCAATACCAACCGCAACG	898
Qy	781	AGTCATGATGAGGACCTCGCAACATCCGGCTC	AGGGTCCGGAGAACGAGAACCCACACA	840
Db	899	AGTCATGATGAGGACCTCGCAACATCCGGCTC	AGGGTCCGGAGAACGAGAACCCACACA	958
Qy	841	ACTTCTGATGCGGTCAACAAGCAAGGGCTTTTC	ATCAACAATGCCACCCACTCCCTCACCA	900
Db	959	ACTTCTGATGCGGTCAACAAGCAAGGGCTTTTC	ATCAACAATGCCACCCACTCCCTCACCA	1018
Qy	901	CAACCAAGCAGCCACAGCAGGAGGAGAAACA	CAACAACCATTTCCCAAGATGCTGTGACT	960
Db	1019	CAACCAAGCAGCCACAGCAGGAGGAGAAACA	CAACAACCATTTCCCAAGATGCTGTGACT	1078
Qy	961	GAACTAGACAAAAATAACAGAACTGACACAC	CGGTGCATGCGCCCTCATATACATCAACA	1020
Db	1079	GAACTAGACAAAAATAACAGAACTGACACAC	CGGTGCATGCGCCCTCATATACATCAACA	1138
Qy	1021	ATCTCTACTATACACACTCCCAACACAACTTC	AGCAGCACTCTCTCTGCACCATTTACAAAAC	1080
Db	1139	ATCTCTACTATACACACTCCCAACACAACTTC	AGCAGCACTCTCTCTGCACCATTTACAAAAC	1198
Qy	1081	ACCACCAATGACAAACACAGAGCACCAATCA	CTGAAAAATGAGCAAAACCACTGCCCTCG	1140
Db	1199	ACCACCAATGACAAACACAGAGCACCAATCA	CTGAAAAATGAGCAAAACCACTGCCCTCG	1258
Qy	1141	ATAACAAACCTTCCTCAACGGGAAATCCCA	CCACAGCAAAAGAGCACAGCAGCAAAAAA	1200
Db	1259	ATAACAAACCTTCCTCAACGGGAAATCCCA	CCACAGCAAAAGAGCACAGCAGCAAAAAA	1318
Qy	1201	GGCCCCGCCACAAACGGCACCAGAAACAGACA	AAATGAGCATTTTCCACAGTCTCTCCCCCACC	1260
Db	1319	GGCCCCGCCACAAACGGCACCAGAAACAGACA	AAATGAGCATTTTCCACAGTCTCTCCCCCACC	1378
Qy	1261	CCCAGCTCGACTGCACAACTCTTTGTATATTT	TCAGAAGAACGATGATTTTCTGAAA	1320
Db	1379	CCCAGCTCGACTGCACAACTCTTTGTATATTT	TCAGAAGAACGATGATTTTCTGAAA	1438
Qy	1321	GAAGGTGATATATTTCCCGTTTTTATAGATGG	GTATATAAATCTGAAATTTGATTTGATCCA	1380
Db	1439	GAAGGTGATATATTTCCCGTTTTTATAGATGG	GTATATAAATCTGAAATTTGATTTGATCCA	1498
Qy	1381	ATCCGAAACACAGAAAACATCTTTTGATGAA	ATCTCCCGAGTCTTAAATCTTCAATGATGAG	1440
Db	1499	ATCCGAAACACAGAAAACATCTTTTGATGAA	ATCTCTAGTTCTTGTGCTCGGTGAGGAA	1558

QY	1441	GAACACACACTCCCGCGAATATACAGTTTAACTTTCTCTTATTTCTCTGATATAAAATGGA	1501
Db	1559	GATCAACATGCCTCCCGCAATATATAGTTTAACTTTATCTTATTTCTCTAATATAAATGAG	1518
QY	1501	GATACTGCCCTACTCTCGGGGAAAAAGAGATGATTTGTGATCGAGAGTTGAGGATTTGGAGT	1560
Db	1619	AACACTGCCCTACTCTGAGAAAAATGAGATGATTTGTGATCGAGATTTAAGATTTTGAGC	1678
QY	1561	GTGCAGGAGACGATTTTGGCGCAGGCGCTTAGCTGGATACCATTTTTTGGCCCTGGAAATC	1620
Db	1679	GTTTCAGGAGGATGACCTGGCGCAGGCGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAAT	1738
QY	1621	GAAGGACTTATCTACTGCGGCTTTAATCAAAAAATCAGAACAAATTTAGTTTGTAGTTTGAGG	1680
Db	1739	GAAGGACTTTTACACTGCTGTTTTTAATTTAAAAAATCAAAAACAATTTGGTCTCGCAGTTGAGG	1798
QY	1681	CGCTTAGCTAATCAAACTGCTAAATCCTTCGGAGCTCTTTGTTAAGGGTCAACAACGAGGAA	1740
Db	1799	CGTCTAGCCAAATCAAACTGCAAAATCCTTGGAACTCTTATTGAGAGTCAACAATGAGGAA	1858
QY	1741	AGGACATTTTCCCTTAATCAATAGGCATSCAAATTTGACTTTTTTGTCTTACGAGTGGGGCGGA	1800
Db	1959	AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTCTCAACAAGATGGGAGGA	1918
QY	1801	ACATGCAAGGTGCTAGACCTGATTTGTTCATAGGAATAGAAATCTATCTTAAAAATATC	1860
Db	1919	ACATGCAAAAGTCTTGGACCTGATTTGTTCATCGGATAGAGACTTTGTCCAAAAATATT	1978
QY	1861	TCAGAACAAATCGACAAATCAGAAAGGATGACAAAGGAGGAGAACTGGCTGGGGTCTA	1920
Db	1979	TCAGAGCAAAATTGACCAAAATTTAAAGAGACGAACAAAGAGGGGACCTGGTGGGGTCTG	2038
QY	1921	GGTGGCAAAATGGTGGACATCTGACTGGGGGTGTTCTTCAACCAATTTGGGCATCTCTACTA	1980
Db	2039	GGTGTGTAATGGTGGACATCCGACTGGGGGTCTCTTACTTAACCTTGGGCATTTTGTCTACTA	2098
QY	1981	TTATCTATAGCTGTTCTTGATGCTCTGCTCTGATCTGTCGTATCTTCACTAAATACATT	2040
Db	2099	TTATCCATAGCTGTCTTGATTTGCTCTATCCTGATTTGTGATTTGTGATCTTTTACTAAATATATC	2158
QY	2041	GGATGA	2046
Db	2159	GGATAA	2164
RESULT 4			
ACC71550			
ID	ACC71550 standard; DNA; 7778 BP.		
XX	ACC71550;		
XX	10-JUL-2003 (first entry)		
DT	10-JUL-2003 (first entry)		
DE	VR6701 (pVR1012-Marburg) plasmid.		
XX	Virucide; vaccine; immune response; Ebola virus; Marburg virus;		
KW	Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;		
KW	viral infection; filovirus; circular; cyclic; ds.		
XX	Synthetic.		
OS	WO2003028632-A2.		
PN	10-APR-2003.		
PD	24-SEP-2002; 2002WO-US030251.		
XX	01-OCT-2001; 2001US-0326476P.		
PR	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
XX	Nabel GJ, Yang Z, Sullivan N, Sanchez A;		
XX			

DR	WPI; 2003-371961/35.	Db	2634	AGTAACGGACGCAACGAATGACACTGGATGTTTTCGGGGCTCTTCAAGATAACAATTCT	2693
XX		Qy	661	ACAAGAACCAAAACATGTGCTCCCTCAAAATACCTCCACACTGCCACAGCCGCTCGG	720
PT	New bimodal priming and boosting compositions, useful as viral vaccines, specifically for eliciting an immune response against a filovirus or a	Dd	2694	ACAAGAACCAAAACATGTGCTCCCTCAAAATAGCTCCACACTGCCACAGCCGCTCGG	2753
PT	disease caused by infection with filovirus.	Qy	721	GAGATCAAACTCACAGCACCCCAACTGATGCGCCACCAAACTCAATACCAACGAGCCCAAGC	780
XX	Claim 1; Page 185-187; 219pp; English.	Dd	2754	GAGATCAAACTCACAGCACCCCAACTGATGCGCCACCAAACTCAATACCAACGAGCCCAAGC	2813
CC	The present invention relates to a bimodal priming composition and	Qy	781	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA	840
CC	boosting composition for priming and boosting an immune response to an	Dd	2814	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA	2873
CC	antigen in an individual. The compositions comprise (a) a priming	Qy	841	ACTTCTGATGGGTACCAAGCAAGGGTTTCATCAACAATGCCACCCACTCCCTCACA	900
CC	composition comprised of a DNA plasmid comprising a nucleic acid molecule	Dd	2874	ACTTCTGATGGGTACCAAGCAAGGGTTTCATCAACAATGCCACCCACTCCCTCACA	2933
CC	encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza	Qy	901	CAACCAAGCAGCCACACAGCAGGAGGAAACAACAACCAATTCCTCCAGATGCTGTGACT	960
CC	virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a	Dd	2934	CAACCAAGCAGCCACACAGCAGGAGGAAACAACAACCAATTCCTCCAGATGCTGTGACT	2993
CC	DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting	Qy	961	GAATAGACAAATAACACAACTGCACAACTGCCACCGTCCCTCATATAACTACACACA	1020
CC	composition comprised of a replication-deficient adenovirus, comprising a	Dd	2994	GAATAGACAAATAACACAACTGCACAACTGCCACCGTCCCTCATATAACTACACACA	3053
CC	nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,	Qy	1021	ATCTCTACTAAACACACTCCAAACACAACTTCAGCAGCTCTCTCTGCAACCAATTAACAAC	1080
CC	paramyxovirus, or influenza virus glycoprotein or nucleoprotein or	Dd	3054	ATCTCTACTAAACACACTCCAAACACAACTTCAGCAGCTCTCTCTGCAACCAATTAACAAC	3113
CC	epitope-bearing domain, or a replication deficient adenovirus selected	Qy	1081	ACCACTATGACACACACAGCAGCACAATCACTGAAATGAGCAACCAAGTCCGCCCCCTCG	1140
CC	from constructs Acc71521-ACC71563. The compositions are useful as viral	Dd	3114	ACCACTATGACACACAGCAGCACAATCACTGAAATGAGCAACCAAGTCCGCCCCCTCG	3173
CC	vaccines, specifically for eliciting an immune response against a	Qy	1141	ATAACAACCTCGCTCCAAACGGGAAATCCCAACAGCAAGAGAGCAGCAGCAAGAAAAA	1200
CC	filovirus or a disease caused by infection with filovirus e.g. Ebola	Dd	3174	ATAACAACCTCGCTCCAAACGGGAAATCCCAACAGCAAGAGAGCAGCAGCAAGAAAAA	3233
CC	virus infection. The present sequence has a Marburg glycoprotein (GP)	Qy	1201	GGCCCCGCCACCAACGGCAGCCAAACACAGCAAAATGAGCAATTTTCCAGTCTCTCCCCCACC	1260
CC	open reading frame, Musoke strain, cloned into a VRC8700 backbone	Dd	3234	GGCCCCGCCACCAACGGCAGCCAAACACAGCAAAATGAGCAATTTTCCAGTCTCTCCCCCACC	3293
XX	Sequence 7778 BP; 2148 A; 1892 C; 1719 G; 2019 T; 0 U; 0 Other;	Qy	1261	CCAGCTCGACTGCACAACATCTTTGTATATTTTCAGAGAAAGCGATCCATTTCTCGGAA	1320
SQ	Query Match 90.0%; Score 1841.2; DB 10; Length 7778;	Dd	3294	CCAGCTCGACTGCACAACATCTTTGTATATTTTCAGAGAAAGCGAAGTATCCTCTGAGG	3353
	Best Local Similarity 93.7%; Pred. No. 0;	Qy	1321	GAAGTGATATATTTCCCGTTTTAGATGGGTTAATAATACTGAAATTTGATTTTGATCCA	1380
	Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;	Dd	3354	GAAGTGATATATTTCCCGTTTTAGATGGGTTAATAATACTGAAATTTGATTTTGATCCA	3413
Qy	1 ATGAAGACACATGTTTCTTATAGCTTTATCTTAAATTCAGGGACAAAAATCTCCC	60			
Dd	2034 ATGAAGACACATGTTTCTTATAGCTTTATCTTAAATTCAGGGACAAAAATCTCCC	2093			
Qy	61 ATTTAGATAGCTAGTAAATCAACCCCAAAATGCGATGCTCCGGAAT	120			
Dd	2094 ATTTAGATAGCTAGTAAATCAACCCCAAAATGCGATGCTCCGGAAT	2153			
Qy	121 CTCAGAGACAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT	180			
Dd	2154 CTCAGAGACAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAGTTGCT	2213			
Qy	181 GATTCCCCCTTTGGAGCATCCAAGCGATGGGCTTTTCCAGACAGGTGTACCTCCCAAGAT	240			
Dd	2214 GATTCCCCCTTTGGAGCATCCAAGCGATGGGCTTTTCCAGACAGGTGTACCTCCCAAGAT	2273			
Qy	241 GTTGAGTACAGAGGGGGAGGAGCCAAACATGCTACATATAGTGAACGATCCC	300			
Dd	2274 GTTGAGTACAGAGGGGGAGGAGCCAAACATGCTACATATAGTGAACGATCCC	2333			
Qy	301 TCTGGAAATCCTTGCTGTAGATCCTCTCAACATCCGTGACTATCCGAAATGCAAA	360			
Dd	2334 TCTGGAAATCCTTGCTGTAGATCCTCTCAACATCCGTGACTATCCGAAATGCAAA	2393			
Qy	361 ACTATCCATCATTAATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	420			
Dd	2394 ACTATCCATCATTAATCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	2453			
Qy	421 GGAGCATTTTCTGTATGATCGCATTGCTCCACACATGATACCGAGGCAAGTCTTC	480			
Dd	2454 GGAGCATTTTCTGTATGATCGCATTGCTCCACACATGATACCGAGGCAAGTCTTC	2513			
Qy	481 ACTGAAGGGAACATAGCAGCTATGTTGCAATAAGACAGTGCACAAATGATTTTCTCG	540			
Dd	2514 ACTGAAGGGAACATAGCAGCTATGTTGCAATAAGACAGTGCACAAATGATTTTCTCG	2573			
Qy	541 CGGCAAGGACAAGGGTACCGTCATATGAATCTGACTTCTACTTAATAATTTGGCAAGT	600			
Dd	2574 CGGCAAGGACAAGGGTACCGTCATATGAATCTGACTTCTACTTAATAATTTGGCAAGT	2633			
Qy	601 AGTAACGGACCGCAACGATCACACTGATGTTTTCGGGGCTCTTCAAGATACAAATCT	660			

```
QY 1741 AGGCATTTCTCTTAATCAATAGCAGTCGCAATTCGACTTTTGGCTTACGAGTGGGGCGGA 1800
D 3774 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCAAGAATGGGGAGGA 3833
QY 1801 ACATGCAAGTGGCTAGGACCTGATTGTCATAGGAATAGAAATCTATCTTAAAAATATC 1860
D 3834 ACATGCAAGTGGCTAGGACCTGATTGTCATAGGAATAGAAATCTATCTTAAAAATATC 1893
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGCTTA 1920
D 3894 TCAGAGCAAAATGACCAAAATTAAGGAGGACCAACAAAGGAGGAGGCTGGTGGGGCTG 3953
QY 1921 GGTGGCAATGGTGGACATCTGACTGGGGTCTTCTCACCATTGGGCAATCTCTACTA 1980
D 3954 GGTGGTAAATGGTGACATCCGACTGGGGTCTTCTTACTAACTTGGGCATTTTCTACTA 4013
QY 1981 TTATCTATAGCTGTCTGATTGCTCTGTCCTGATATCTGCTGATCTTCACTAAATACATT 2040
D 4014 TTATCCATAGCTGTCTGATTGCTCTATCCGTATTTGCTGATATCTTACTAAATATATC 4073
QY 2041 GGATGA 2046
D 4074 GGATAA 4079

RESULT 5
ID AA287211 standard; cDNA; 11460 BP.
AC AA287211;
XX
DT 15-SEP-2003 (revised)
DT 08-MAY-2000 (first entry)
XX
DE VEE replicon comprising Marburg virus genes.
XX
KW VEE virus replicon; Venezuelan equine encephalitis; Marburg virus; MBGV;
KW filovirus; Glycoprotein; GP; nucleoprotein; NP; structural protein; VP40;
KW VP35; VP30; VP24; genetic vaccine; antigen delivery;
KW Marburg haemorrhagic fever; cyclic; circular; ss.
XX
OS Marburg virus; str. Musoke.
OS Venezuelan equine encephalitis virus.
OS Chimeric.
XX
FH Key
FT CDS Location/Qualifiers
FT CDS 104..2182
FT CDS /product= "Marburg virus nucleoprotein (NP, AAY77128)"
FT CDS /tag= a
FT CDS 2944..3933
FT CDS /tag= b
FT CDS /product= "Marburg virus structural protein VP35
FT CDS (AAY77130)"
FT CDS 4567..5478
FT CDS /tag= c
FT CDS /product= "Marburg virus structural protein VP40
FT CDS (AAY77129)"
FT CDS 5940..7985
FT CDS /tag= d
FT CDS /product= "Marburg virus glycoprotein (GP, AAY77127)"
FT CDS 8864..9697
FT CDS /tag= e
FT CDS /product= "Marburg virus structural protein VP30
FT CDS (AAY77131)"
FT CDS 10200..10961
FT CDS /tag= f
FT CDS /product= "Marburg virus structural protein VP24
FT CDS (AAY77132)"
XX
PN WO20000616-A2.
XX
PD 06-JAN-2000.
```

```
XX 21-JUN-1999; 99WO-US014174.
PF
XX 29-JUN-1998; 98US-0091403P.
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
PI Hevey MC, Negley DL, Pushko P, Smith JF, Schmaljohn AL;
XX
XX WPI; 2000-160676/14.
DR P-PSDB; AAY77127, AAY77128, AAY77129, AAY77130, AAY77131, AAY77132,
DR AAY77133.
XX
PT Novel Marburg virus vaccines used to induce an immune response against
PT the infection in nonhuman primates.
XX
PS Claim 1; Page 56-57; 57pp; English.
XX
CC The invention relates to novel Marburg virus (MBGV) vaccines. The vaccine
CC of the invention comprises a Venezuelan equine encephalitis virus (VEE)
CC replicon containing copies of the genes encoding Marburg virus (Musoke
CC strain) glycoprotein (GP), a deletion mutant GP (GP-delta-TM),
CC nucleoprotein (NP), and structural proteins VP40, VP35, VP30 and VP24.
CC The Marburg virus genes replace the VEE virus structural protein genes;
CC the result is a self-replicating RNA molecule that encodes its own
CC replicase and transcriptase functions, and in addition makes abundant
CC quantities of the Marburg virus proteins. When replicon RNA is
CC transfected into eukaryotic cells, along with two helper RNAs that
CC express the VEE virus structural proteins, the replicon RNA is packaged
CC into VEE virus-like particles by the VEE virus structural proteins, which
CC are provided in trans. Since the helper RNAs lack packaging signals
CC necessary for further propagation, the resulting VEE replicon particles
CC (VRPs) which are produced are infectious for one cycle but are defective
CC thereafter. On VRP infection of a cell, an abortive infection occurs
CC whereby the cell produces the Marburg virus proteins, is ultimately
CC killed by the infection, but does not produce any viral progeny. The VEE
CC replicon provides a potent tool for vaccination with Marburg virus
CC antigens. The replicons, vectors and constructs are used to produce
CC vaccines against Marburg virus (MBGV) infection (Marburg haemorrhagic
CC fever) in mammals, to elicit immune responses against Marburg antigens,
CC to confer protective immunity, and to reduce disease symptoms and reduce
CC the severity of disease. Studies of non-human primates vaccinated with
CC the replicon indicate that the vaccine will be efficient in protecting
CC humans against Marburg virus. Prior art Marburg virus vaccination
CC strategies have used formalin-inactivated Marburg virus, which are only
CC partially successful at protecting against Marburg virus infection. The
CC present sequence represents a cDNA corresponding to a VEE virus replicon
CC encoding Marburg virus GP, GP-delta-TM, NP, VP40, VP35, VP30 and VP24.
CC (Updated on 15-SEP-2003 to standardise OS field)
XX
SQ Sequence 11460 BP; 3674 A; 2382 C; 2145 G; 3259 T; 0 U; 0 Other;

Query Match 90.0%; Score 1841.2; DB 3; Length 11460;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 1 ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAAATCTCCCC 60
D 5940 ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAAATCTCCCC 5999

QY 61 ATTTAGAGATAGCTAGTAATATCAACCCCAAAATGCGATTCGGTATGCTCCGGAAT 120
D 6000 ATTTAGAGATAGCTAGTAATATCAACCCCAAAATGCGATTCGGTATGCTCCGGAAT 6059

QY 121 CTCAGAGACAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGGCAAAAGTTGCT 180
D 6060 CTCAGAGACAGAGACGCTCCATCTGATGGGATTCACACTGAGTGGGGCAAAAGTTGCT 6119

QY 181 GATTCCCTTTGGAGGCATCAAGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
D 6120 GATTCCCTTTGGAGGCATCAAGCGATGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 6179

QY 241 GTTGAGTACACAGAGGGGGAGGAAGCAAAACATGCTACATATATAAGTGTAAACGGATCCC 300
```


PN WO200008131-A2.
XX 17-FEB-2000.
XX 04-AUG-1999; 99WO-US017702.
XX 04-AUG-1998; 98US-0095242P.
PR 15-DEC-1998; 98US-0112405P.
XX (PURD) PURDUE RES FOUND.
XX PA
XX Sanders DA, Kuhn RJ, Jeffers SA, Sharkey CM, North CL;
PI Fischbach MA;
XX WPI; 2000-224030/19.
DR P-PSDB; AAY70075.
XX
DR Cells that produce inventive pseudotyped retroviruses having a broad host
PT range useful for introducing nucleotide sequences into target cells.
XX
XX Example 11; Page 60-61; 65pp; English.
XX
XX The present sequence encodes the Marburg virus envelope glycoprotein.
CC Plasmid pMBGpi was produced by cloning into the plasmid pSP72 nucleotide
CC sequences corresponding to nucleotides 5931-8033 from the Marburg virus
CC genome. This is used for generation of stable cell lines transiently
CC producing Marburg-MMLV (Moloney murine leukemia virus) pseudotyped
CC retrovirus with a broad host range. These cells having different viral
CC glycoproteins in its lipid bilayer with a pseudotyped retrovirus is
CC transduced with a desired ribonucleotide sequence. This is used to
CC identify screening agents effective in blocking viral entry into a cell.
CC These agents may be immunological agents like monoclonal or polyclonal
CC antibodies. The pharmacological agents include proteins, peptides or
CC various chemical agents. The pseudotyped retrovirus may be useful in
CC methods of identifying cell surface receptors that allow viral entry
XX
XX Sequence 2100 BP; 670 A; 532 C; 408 G; 490 T; 0 U; 0 Other;
SQ

Query Match 89.2%; Score 1825.2; DB 3; Length 2100;
Best Local Similarity 93.6%; Pred. No. 0;
Matches 1915; Conservative 0; Mismatches 128; Indels 3; Gaps 1;
1 ATGAAGACCATGTTTCTTATCAGTCTTATCTTAAATTCAGGGACAAAAATCTCCC 60
10 ATGAAGACCATGTTTCTTATCAGTCTTATCTTAAATTCAGGGACAAAAATCTCCC 69
61 ATTTTATGAGATGCTAGTAAATCAACCCCAAAATGTTGATGCTCCGGAACT 120
70 ATTTTATGAGATGCTAGTAAATCAACCCCAAAATGTTGATGCTCCGGAACT 129
121 CTCAGAGACAGAGACGTCCTATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
130 CTCAGAGACAGAGACGTCCTATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 189
181 GATTCCTCTTGGAGGATCCAGCGATGGGTTTCAGGACAGGTGATCCTCCCAAGAT 240
190 GATTCCTCTTGGAGGATCCAGCGATGGGTTTCAGGACAGGTGATCCTCCCAAGAT 249
241 GTTGAATACAGAGGGGGAGGAGCCAAACATGCTCAATATAGTAAAGTAAAGGATCCC 300
250 GTTGAATACAGAGGGGGAGGAGCCAAACATGCTCAATATAGTAAAGTAAAGGATCCC 309
301 TCTGGAATATCTCTGTTAGATCCTCTACCAACATCCGTGATCTCCGAAATGCAAA 360
310 TCTGGAATATCTCTGTTAGATCCTCTACCAACATCCGTGATCTCTTAATGCAAA 369
361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCTTCATTTATGG 420
370 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCTTCATTTATGG 429
421 GGAGCATTTTCTGATGATGCTGCTCCCAACATGATGATGATGATGATGATGATGAT 480
430 GGAGCATTTTCTGATGATGCTGCTCCCAACATGATGATGATGATGATGATGATGAT 489

QY 481 ACTCAAGGGAAACATAGCAGCTATGATTGTCAATTAAGACAGTGCACAAAATGATTTTCTCG 540
DB 490 ---GNAAGGGAAACATAGCAGCTATGATTGTCAATTAAGACAGTGCACAAAATGATTTTCTCG 546
QY 541 CGGCAAGGACAAAGGGTACCGTTCATATGAATCTGACTTCTACTAATAATATTTGGACAAGT 600
DB 547 CGGCAAGGACAAAGGGTACCGTTCATATGAATCTGACTTCTACTAATAATATTTGGACAAGT 606
QY 601 AGTAAACGGAAACGCAAAACGAATGACACATGATGTTTTCGGCGCTCTTCAAGAAATACAATTTCT 660
DB 607 AGTAAACGGAAACGCAAAACGAATGACACATGATGTTTTCGGCGCTCTTCAAGAAATACAATTTCT 666
QY 661 ACAAAGAACCAAAACATGTTCTCCGTCGAAATATCTCCACACATCTGCCCCACAGCCGCTCG 720
DB 667 ACAAAGAACCAAAACATGTTCTCCGTCGAAATATCTCCACACATCTGCCCCACAGCCGCTCG 726
QY 721 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780
DB 727 GAGATCAAACTCACAAAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 786
QY 781 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
DB 787 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 846
QY 841 ACTTCTGATGCGGTACCAAGCAAGGGTTTTCATCAACAATGCGCAACCCCTCCCTCAACA 900
DB 847 ACTTCTGATGCGGTACCAAGCAAGGGTTTTCATCAACAATGCGCAACCCCTCCCTCAACA 906
QY 901 CAACCAAGCAGCCACAGCAGGAGGAAACAAACCAACCAATTCCTCAGATGCTGTGACT 960
DB 907 CAACCAAGCAGCCACAGCAGGAGGAAACAAACCAACCAATTCCTCAGATGCTGTGACT 966
QY 961 GAACCTAGACAAAAATAACACAACTGCAACCGTCCATGCCCCCTCATAACTACCAACA 1020
DB 967 GAACCTAGACAAAAATAACACAACTGCAACCGTCCATGCCCCCTCATAACTACCAACA 1026
QY 1021 ATCTCTACTAAACAACCTCCAAACAACAATTTAGAGCACTCTCTCTGCAACCAATTAACAAC 1080
DB 1027 ATCTCTACTAAACAACCTCCAAACAACAATTTAGAGCACTCTCTCTGCAACCAATTAACAAC 1086
QY 1081 ACCACCAATGACACACAGCAGCACAATCACTGAAATGAGCAACCAAGTCCGCCCTCG 1140
DB 1087 ACCACCAATGACACACAGCAGCACAATCACTGAAATGAGCAACCAAGTCCGCCCTCG 1146
QY 1141 ATAAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAGAGACCAAGCAGCAAAAAA 1200
DB 1147 ATAAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAGAGACCAAGCAGCAAAAAA 1206
QY 1201 GGCCCCGCCCAACAGCGCAACCAACAAGCAATGAGCAATTCACAGTCTCTCCGCCACC 1260
DB 1207 GGCCCCGCCCAACAGCGCAACCAACAAGCAATGAGCAATTCACAGTCTCTCCGCCACC 1266
QY 1261 CCGAGCTGCACTGCACACATCTTGTATATTTTCAGAGAAAGCGATCGATTTTCTGGAA 1320
DB 1267 CCGAGCTGCACTGCACACATCTTGTATATTTTCAGAGAAAGCGATCGATTTTCTGGAG 1326
QY 1321 GAAGTGATATATTTCCCGTTTATAGTGGGTTAATAATACTGAAATTTGATTTGATCCA 1380
DB 1327 GAAGTGATATATTTCCCGTTTATAGTGGGTTAATAATACTGAAATTTGATTTGACCCA 1386
QY 1381 ATCCCAACACAGAAACAATCTTTGATGAATCTCCAGCTTTTAATTAATCTCAACTAATGAG 1440
DB 1387 GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTCTCTGGTCCCTCGGCTGAGGAA 1446
QY 1441 GAACAAACACATCCCGCAATATCAGTTTAACTTTCTTATTTCTGTATTAATAATGGA 1500
DB 1447 GATCAACATGCTCCCGCAATATTAAGTTAACTTTTATCTTATTTCTTAATAATAATGAG 1506
QY 1501 GATCTGCTACTCTGGGGAAAAACGAGATGATGATGATGATGATGATGATGATGATGATGAGT 1560
DB 1507 AACCTGCTACTCTGGAGAAAAATGAGATGATGATGATGATGATGATGATGATGATGATGAGC 1566

QY 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 120
Db 2094 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 2153
QY 121 CTCAGAGACAGAGAGCGTCCATCTGATGGGATTTCAACATGAGTGGGCGAAAGTGTCT 180
Db 2154 CTCAGAGACAGAGAGCGTCCATCTGATGGGATTTCAACATGAGTGGGCGAAAGTGTCT 2213
QY 181 GATTCCTCTTTGGAGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 240
Db 2214 GATTCCTCTTTGGAGGCATCCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAAT 2273
QY 241 GTTAGATACAGAGGGGGAGAGCAAAACATGCTACAATATAGTGTAAACGATCCC 300
Db 2274 GTTAGATACAGAGGGGGAGAGCAAAACATGCTACAATATAGTGTAAACGATCCC 2333
QY 301 TCTGGAAATCCTTCTGTAGATCCTCTACCAACATCCGTGACTATCCGNAATGCAAA 360
Db 2334 TCTGGAAATCCTTCTGTAGATCCTCTACCAACATCCGTGACTATCCGNAATGCAAA 2393
QY 361 ACTATCCCATCATATTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATATGG 420
Db 2394 ACTATCCCATCATATTCAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATATGG 2453
QY 421 GGAGCATTTTCTGTATGATCGCATTTGCTCCACAACAATGTACCGAGGCAAACTCTTC 480
Db 2454 GGAGCATTTTCTGTATGATCGCATTTGCTCCACAACAATGTACCGAGGCAAACTCTTC 2513
QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATNAGACAGTGCACAAAATGATTTTCTCG 540
Db 2514 ACTGAAGGGAACATAGCAGCTATGATTTGTCATNAGACAGTGCACAAAATGATTTTCTCG 2573
QY 541 CGGCAAGGACAGAGGTACCGTCAATATGATATCTGACTTCTACTATAAATATTGGACAAGT 600
Db 2574 CGGCAAGGACAGAGGTACCGTCAATATGATATCTGACTTCTACTATAAATATTGGACAAGT 2633
QY 601 ACTAACGGAAACGCAACGAAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAAATCT 660
Db 2634 ACTAACGGAAACGCAACGAAATGACACTGGATGTTTCGGCGCTCTTCAAGAAATACAAATCT 2693
QY 661 ACAAGAACCAAAACATGTGCTCGGTCCAAATATACCTTCACGACTGCCACAGCCCGTCCG 720
Db 2694 ACAAGAACCAAAACATGTGCTCGGTCCAAATATACCTTCACGACTGCCACAGCCCGTCCG 2753
QY 721 GAGATCAAACTCAAGACACCCCAACTGATGCCACCAAACTCAATACCAAGGACCCCAAGC 780
Db 2754 GAGATCAAACTCAAGACACCCCAACTGATGCCACCAAACTCAATACCAAGGACCCCAAGC 2813
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGCTCGGAGAAACGAAACCCCAACA 840
Db 2814 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGCTCGGAGAAACGAAACCCCAACA 2873
QY 841 ACTTCTGATCGGTCAACAGCAAGGGCTTTTCATCAACATGCCACCCACTCCCTCAGCA 900
Db 2874 ACTTCTGATCGGTCAACAGCAAGGGCTTTTCATCAACATGCCACCCACTCCCTCAGCA 2933
QY 901 CAACCAAGCACGCCACAGACAGAGGAAACAACAACATTCCTCAAGATGCTGTGACT 960
Db 2934 CAACCAAGCACGCCACAGAGAGAGGAAACAACAACATTCCTCAAGATGCTGTGACT 2993
QY 961 GAACTAGACAAAATAACACAACTGCAACCGTCCATGCCCTCCCTCATAAACAATACCA 1020
Db 2994 GAACTAGACAAAATAACACAACTGCAACCGTCCATGCCCTCCCTCATAAACAATACCA 3053
QY 1021 ATCTCTACTAACACACCTCCAAACAACAATTCAGGACTCTCTCTGCAACCATTAACA 1080
Db 3054 ATCTCTACTAACACACCTCCAAACAACAATTCAGGACTCTCTCTGCAACCATTAACA 3113
QY 1081 ACCACCAATGACACACAGAGCACAATCTGAAAAATGACAAACAGTGCCTCCCTCG 1140
Db 3114 ACCACCAATGACACACAGAGCACAATCTGAAAAATGACAAACAGTGCCTCCCTCG 3173
QY 1141 ATAAACAACCTCGCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGCAAAAAA 1200

Db 3174 ATAAACAACCTCGCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGCAAAAAA 3233
QY 1201 GGCCCGGCACACAGCGCACCAACACAGCAAAATGAGCAATTTCCACAGTCTCTCCCCACCC 1260
Db 3234 GGCCCGGCACACAGCGCACCAACACAGCAAAATGAGCAATTTCCACAGTCTCTCCCCACCC 3293
QY 1261 CCCAGCTCGACTGCAACAAACATCTTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAAA 1320
Db 3294 CCCAGCTCGACTGCAACAAACATCTTGTATATTTTCAGAGAAAGCGAAGTATCTCTCTGGAGG 3353
QY 1321 GAAAGTGATATATTTCCCGTTTTTATAGTGGGTATTAATACTAGAAATTTGATTTTGATCCA 1380
Db 3354 GAAAGCGCATGTTTCCCTTTTCTGGATGGGTAAATAATGCTCCAATTTGATTTTGACCCA 3413
QY 1381 ATCCCAACACAGAGAAACAACTTTGTATGTAATCTCCAGCTTTAATACCTCAATTAATGAG 1440
Db 3414 GTTCCAAATACAAAAACAACTTTGTATGTAATCTCTAGTCTGTGTCCTCGCTCAGGAGAA 3473
QY 1441 GAAACAACACTCCCCCGAAATATCAGTTTAACTTTCTCTTATTTTCTGTATAAAAAATGGA 1500
Db 3474 GATCAACATGCTCTCCCAATATTTAGTTTAACTTTATCTTATTTCTTAATATAATGAG 3533
QY 1501 GATACTGCTACTCTGGGGAAAAACAGATGATTTGTGATGAGAGTTGAGGATTTGGAGT 1560
Db 3534 AACACTGCTACTCTGGGAGAAAAATGAGAAATGATTTGTGATGAGAGTTAAGAAATTTGGAGC 3593
QY 1561 GTGAGAGAGACGATTTTGGCGGAGGCTTAGCTGGATACCAATTTTGGCCCTCGAAATC 1620
Db 3594 GTTCAGAGAGGATGACTGGCCGCGAGGGCTCAGTTGGATACCGTTTTTGGCCCTGGAAT 3653
QY 1621 GAAGGACTCTATACTGCGCGTTTAAATCAAAAATCAGAAACAATTTAGTTTGTAGTTGAGG 1680
Db 3654 GAAGGACTTTACACTGCTGTTTTTAAATTAATAATCAAAAACAATTTGGTCTGCAGGTTGAGG 3713
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTTTGGAGCTCTTTGTAAGGCTCAACAACCGAGAA 1740
Db 3714 CGCTTAGCCAATCAAACTGCCAAATCTTGGAACTCTTTATTTGAGAGTCAACAACCTGAGGAA 3773
QY 1741 AGGACATTTTCTTAATCAATAGCATGCAATTCGACTTTTGTCTTACGAGGTGGCGCGGA 1800
Db 3774 AGAATTTCTCTTAATCAATAGCATGCTATTGACTTTCTACTCAAGATGGGAGAGA 3833
QY 1801 ACATGCAAGGTGCTAGGACCTGATTGTTCATAGAGATAGAGATCTATCTAAAAATATC 1860
Db 3834 ACATGCAAGGTGCTTGGACCTGATTGTTCATCGGATAGAAGACTTGTCCAAAAATATT 3893
QY 1861 TCAGAACAAATCGACAAATCAGAAAGATGAAACAAAAGGAGAAACTGGCTGGGCTCTA 1920
Db 3894 TCAGAGCAAAATTTGACCAAAATTAATAAGGACGAAACAAAAGAGGAGGACTGGTTGGGGTCTG 3953
QY 1921 GGTGCAAAATGGTGGACATCTGACTGGGGT 1950
Db 3954 GGTGTAATGGTGGACATCCGACTGGGGT 3983
RESULT 9
ACCT71552
ID ACCT71552 standard; DNA; 8256 BP.
XX
AC ACCT71552;
XX
DT 10-JUL-2003 (first entry)
XX
DE VRC6710 (pAdapt Marburg GP (dTM)) plasmid.
XX
KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
viral infection; filovirus; circular; cyclic; ds.
XX
OS Synthetic.
XX
PN W02003028632-A2.

```
XX PD 10-APR-2003.
XX PF 24-SEP-2002; 2002MO-US030251.
XX PR 01-OCT-2001; 2001US-0326476P.
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX DR WPI; 2003-371961/35.
XX PT New binodal priming and boosting compositions, useful as viral vaccines,
XX PR specifically for eliciting an immune response against a filovirus or a
XX PR disease caused by infection with filovirus.
XX PS Claim 1; Page 190-193; 219pp; English.
XX CC The present invention relates to a binodal priming composition and
XX CC boosting composition for priming and boosting an immune response to an
XX CC antigen in an individual. The compositions comprise (a) a priming
XX CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
XX CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
XX CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
XX CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
XX CC composition comprised of a replication-deficient adenovirus, comprising a
XX CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
XX CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
XX CC epitope-bearing domain, or a replication deficient adenovirus selected
XX CC from constructs ACC71521-ACC71563. The compositions are useful as viral
XX CC vaccines, specifically for eliciting an immune response against a
XX CC filovirus or a disease caused by infection with filovirus e.g. Ebola
XX CC virus infection. The present sequence is an adenovirus shuttle vector,
XX CC expressing the Marburg virus glycoprotein without its transmembrane and
XX CC intracellular domains
XX SQ Sequence 8256 BP; 2048 A; 2049 C; 2108 G; 2051 T; 0 U; 0 Other;
Query Match 86.4%; Score 1767.6; DB 10; Length 8256;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 1036; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
QY 1 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATCTTAATCTAAGGACAAAAATCTCCC 60
DB 1431 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATCTAAGGACAAAAATCTCCC 1490
QY 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGGATTCGGTATGCTCCGGAACT 120
DB 1491 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGGATTCGGTATGCTCCGGAACT 1550
QY 121 CTCCAGAGACAGAGAGAGCTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
DB 1551 CTCCAGAGACAGAGAGAGCTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 1610
QY 181 GATTCCTTTGGAGGCAATCCAAAGCGATGGGCTTTTACAGACAGGTTGTAACCTCCCAAGAT 240
DB 1611 GATTCCTTTGGAGGCAATCCAAAGCGATGGGCTTTTACAGACAGGTTGTAACCTCCCAAGAT 1670
QY 241 GTTGGATACAGAGGGGGAGAGCCAAACATGCTACATATAGTGAACGATCCC 300
DB 1671 GTTGGATACAGAGGGGGAGAGCCAAACATGCTACATATAGTGAACGATCCC 1730
QY 301 TCTGGAAAATCTTCTGTTAGATCCTCTACCAACATCCCGTACTATCCGAAAATGCAAA 360
DB 1731 TCTGGAAAATCTTCTGTTAGATCCTCTACCAACATCCCGTACTATCCGAAAATGCAAA 1790
QY 361 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
DB 1791 ACTATCCATCATATTAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 1850
QY 421 GGAGCATTTTCTGTATGATCGATTCGCTCCACAAATGATACGAGGCAAGTCTTC 480
DB 1851 GGAGCATTTTCTGTATGATCGATTCGCTCCACAAATGATACGAGGCAAGTCTTC 1910
QY 481 ACTGAAGGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTCTCG 540
DB 1911 ACTGAAGGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTCTCG 1970
QY 541 CGGCAAGGACAAAGGGTACCGTTCATATGATCTGATCTTCTACTATAAATATTTGGACAAGT 600
DB 1971 CGGCAAGGACAAAGGGTACCGTTCATATGATCTGATCTTCTACTATAAATATTTGGACAAGT 2030
QY 601 AGTAAACGGAAACGCAACGAATGACACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAATTTCT 660
DB 2031 AGTAAACGGAAACGCAACGAATGACACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAATTTCT 2090
QY 661 ACAAGAAACCAAAATGCTGCTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 720
DB 2091 ACAAGAAACCAAAATGCTGCTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2150
QY 721 GAGATCAAACTCACAGCACCCCAACTGATGTCACCAAACTCAATACCAACGACCAAGC 780
DB 2151 GAGATCAAACTCACAGCACCCCAACTGATGTCACCAAACTCAATACCAACGACCAAGC 2210
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA 840
DB 2211 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA 2270
QY 841 ACTTCTGATGGGTTCACCAAGCAAGGGCTTTTCATCAACAATGCGACCCACTCCCTCACA 900
DB 2271 ACTTCTGATGGGTTCACCAAGCAAGGGCTTTTCATCAACAATGCGACCCACTCCCTCACA 2330
QY 901 CAACCAAGCAGCCACAGCAGGAGGAGAAACAACAACCAATTCCTCAAGATGCTGTGACT 960
DB 2331 CAACCAAGCAGCCACAGCAGGAGGAGAAACAACAACCAATTCCTCAAGATGCTGTGACT 2390
QY 961 GAACCTAGACAAAAATAACAACTGTCACAAACCGTCCATGCGCCCTCTATPAAACACTACCACA 1020
DB 2391 GAACCTAGACAAAAATAACAACTGTCACAAACCGTCCATGCGCCCTCTATPAAACACTACCACA 2450
QY 1021 ATCTCTACTAAACAACCTCCAAACAACAATTCAGCACTCTCTCTGCAACCAATTAACAAAAC 1080
DB 2451 ATCTCTACTAAACAACCTCCAAACAACAATTCAGCACTCTCTCTGCAACCAATTAACAAAAC 2510
QY 1081 ACACCAATGACACACAGCAGCAGCAATCACTCAAAATGAGCAAAACAGTGCCTCCCTCG 1140
DB 2511 ACACCAATGACACACAGCAGCAGCAATCACTCAAAATGAGCAAAACAGTGCCTCCCTCG 2570
QY 1141 ATAAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGCAGCAAAAAA 1200
DB 2571 ATAAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGCAGCAAAAAA 2630
QY 1201 GGCCCCGCCACACAGGCAACCAACACGCAAAATGAGCAATTTTCAAGTCTCTCCCTCCACC 1260
DB 2631 GGCCCCGCCACACAGGCAACCAACACGCAAAATGAGCAATTTTCAAGTCTCTCCCTCCACC 2690
QY 1261 CCGAGCTGCTGACACACATCTTGTATATTTTCAAGAAAGAGCGATCGATTTCTTGGAAA 1320
DB 2691 CCGAGCTGCTGACACACATCTTGTATATTTTCAAGAAAGAGCGATCGATTTCTTGGAAA 2750
QY 1321 GAAGTGATATATTTCCCGTTTTAGATGGGTAAATAAATACTGAAATTTGATTTTGCATCCA 1380
DB 2751 GAAGTGATATATTTCCCGTTTTAGATGGGTAAATAAATACTGAAATTTGATTTTGCATCCA 2810
QY 1381 ATCCCAAAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTTAAATCTTCAACTAATGAG 1440
DB 2811 GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTCTGGTGGCTCGCTCGGAGAA 2870
QY 1441 GAAACAACACTCCCGCGAATATCAGTTTAACTTTCTTCTTATTTCTCTGATAAAAAATGGA 1500
DB 2871 GATCAACATGCTCCCGCGAATATCAGTTTAACTTTTAACTTTTCTTCTTATTAATAATGAG 2930
QY 1501 GATCTGCTACTCTGGGGGAAACGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1560
DB 2931 AACACTGCTACTCTGGGAGAAATGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 2990
```

```
QY 1561 GTGAGGAGGACGATTTGGCGCAGGGCTTAGCTGGATACCAATTTTGGCCCTGGAATC 1620
DB 2991 GTTCAGGAGGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTGGCCCTGGAAT 3050
QY 1621 GNAGGACTCTATCTGCGCGGTTTAAATCAAAAATCAGAACAAATTTAGTTGAGTTGAGG 1680
DB 3051 GAAGGACTTTACACTGCTGCTTTTAAATTAATAATCAAAAATTTGCTGTCGAGGTTGAGG 3110
QY 1681 CGCTTAGCTAAATCAAACTGCTAAATCCCTGGAGCTCTTGTAAAGGTCACAAACCCAGGAA 1740
DB 3111 CGTCTAGGCAATCAAACTGCTAAATCCCTGGAGCTCTTGTAAAGGTCACAACTGAGGAA 3170
QY 1741 AGGACATTTCTTAAATCAATAGGATGCAATTTGACTTTTTCCTTTACGAGGTGGGGCGGA 1800
DB 3171 AGAACATTTCTTAAATCAATAGATGCTATTGACTTTTACTCACAAGATGGGGGAGGA 3230
QY 1801 ACATGCAAGGTCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
DB 3231 ACATGCAAGGTCTTGGACCTGATTTGTCATCGGGATAGAGACTTTGTCCAAAAATAT 3290
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAAGGAGGAAACTGGCTGGGGTCTA 1920
DB 3291 TCAGAGCAATTTGACAAATTAATAAGGACGAAACAAAAGAGGGGACTGTTGGGGTCTG 3350
QY 1921 GGTGGCAATGTGTGACATCTGACTGGGT 1950
DB 3351 GGTGGTAAATGTGGACATCCGACTGGGT 3380
```

RESULT 10

ADM48345

ID ADM48345 standard; DNA; 19112 BP.

AC ADM48345;

DT 18-NOV-2004 (first entry)

DE Marburg virus viral protein genomic DNA.

XX Filovirus; viral protein; NP protein; VP30 protein; VP35 protein;
KW VP40 protein; L protein; membrane-associated protein; matrix protein;
KW polymerase complex protein; minor nucleoprotein; glycoprotein; GP;
KW viral mutagenesis study; vaccine; gene therapy; gene; ds.

OS Marburg virus.

```
XX Key Location/Qualifiers
FH CDS 103..2190
FT /*tag= a
FT /product= "NP protein"
FT 2944..3933
FT /*tag= b
FT /product= "VP35 protein"
FT 4567..5478
FT /*tag= c
FT /product= "VP40 protein"
FT 5940..7985
FT /*tag= d
FT /product= "GP protein"
FT 8867..9712
FT /*tag= e
FT /product= "VP30 protein"
FT 10205..10966
FT /*tag= f
FT /product= "VP24 protein"
FT 11479..18474
FT /*tag= g
FT /product= "L protein"
```

US2003215794-A1.

20-NOV-2003.

```
XX 29-JAN-2003; 2003US-00353856.
XX 31-JAN-2002; 2002US-0353972P.
XX (KAWA/) KAWAKA Y.
XX (JASE/) JASENOSKY L D.
XX (NEUM/) NEUMANN G.
XX Kawaoka Y, Jasenovsky LD, Neumann G;
XX WPI; 2004-010776/01.
XX P-PSDB; ADM48338, ADM48339, ADM48340, ADM48341, ADM48342, ADM48343,
XX ADM48344.
XX preparing filovirus useful in producing vaccines and gene therapy
XX vectors, comprises contacting a cell with a vector comprising a promoter
XX operably linked to a filovirus genomic cDNA or protein.
XX Disclosure; SEQ ID NO 27; 104pp; English.
```

XX The invention relates to a method of preparing filovirus. The method involves contacting a cell with a vector comprising a promoter operably linked to a filovirus genomic cDNA or its portion linked to a transcription termination sequence or to a DNA segment encoding a filovirus RNA transcriptase-polymerase, filovirus NP, VP30 or VP35, to yield infectious filovirus. The portion of the cDNA when transcribed yields a RNA capable of being packaged into filovirus virions or capable of being replicated in the presence of filovirus proteins. The method is useful for producing filovirus useful in viral mutagenesis studies and in the production of vaccines and gene therapy vectors. The present sequence is marburg virus (NP; L; polymerase complex-VP35; matrix protein-VP40; glycoprotein-GP; minor nucleoprotein-VP30; membrane-associated protein-VP24) genomic DNA.

SQ Sequence 19112 BP; 6176 A; 3757 C; 3578 G; 5601 T; 0 U; 0 Other;

```
Query Match 82.7%; Score 1692.4; DB 13; Length 19112;
Best Local Similarity 89.2%; Pred. No. 0;
Matches 1925; Conservative 0; Mismatches 221; Indels 0; Gaps 0;

QY 1 ATGAAGACCACATGTTCTTATCAGTCTTATCTTAATCAAGGACAAAAATCTCC 60
DB 5940 ATGAAGACCACATGCTCTTTATCAGTCTTATCTTAATCAAGGATAAAACTCC 5999
QY 61 ATTTAGAGTAGCTAGTAATAATCAACCCAAATGCGATTGCGTATGCTCCGGA 120
DB 6000 ATTTAGAGTAGCTAGTAATAATCAACCCAAATGCGATTGCGTATGCTCCGGA 6059
QY 121 CTCAGAGACAGAGACGTCCTATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTG 180
DB 6060 CTCAGAGACAGAGAGATGTCCTATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTG 6119
QY 191 GATTCCCTTTGGAGGCATCCAGGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT 240
DB 6120 GATTCCCTTTGGAGGCATCCAGGATGGGCTTTCAGGACAGGTGTACCTCCCAAGAT 6179
QY 241 GTTGAGTACAGAGGGGGAGGCCAAACATGCTACAATATAAGTGTACGGATCC 300
DB 6180 GTTGAGTACAGAGGGGGAGGCCAAACATGCTACAATATAAGTGTACGGATCC 6239
QY 301 TCTGAAAATCTTGTCTTAGATCTCTCTACCAACATCCGTAATCCGAAATGCAAA 360
DB 6240 TCTGAAAATCTTGTCTTAGATCTCTCTACCAACATCCGTAATCCGAAATGCAAA 6299
QY 361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATGG 420
DB 6300 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTATGG 6359
QY 421 GGAGCATTTTCTGTATGATCGCATTCCTCCACAAATGATGACGAGGCAAGTCTTC 480
DB 6360 GGAGCATTTTCTGTATGATCGCATTCCTCCACAAATGATGACGAGGCAAGTCTTC 6419
```

QY 481 ACTGAGGGAACATAGCAGCTATGATTCTCAATAGACAGATGCGCAAAAATGATTTCCTCG 540
Db 6420 ACTGAAGGGAACATAGCAGCTATGATTCTCAATAGACAGATGCGCAAAAATGATTTCCTCG 6479
QY 541 CGGCAAGGACAGAGGTACCGTCAATGAACTCTGACTTCTACTAATAAATATTGGACAAGT 600
Db 6480 AGGCAAGGACAGAGGTACCGTCAATGAACTCTGACTTCTACTAATAAATATTGGACAAGT 6539
QY 601 AGTAAACGGAACGCAAAACGAATGACACTGGATGTTTCGGGCTCTTCAAGAATATCAAAATCT 660
Db 6540 AACAAATGGAACACAAACGAATGACACTGGATGTTTCGGGCTCTTCAAGAATATCAAAATCT 6599
QY 661 ACAAGACACCAACATGCTCCGTCGCAAAATACCTCCACCACTGCGCCACAGCCGCTCG 720
Db 6600 ACGAAGATCAAAACATGCTCCGTCGCAAAATACCTCCACCACTGCGCCACAGCCGCTCG 6659
QY 721 GAGATCAAACTCACAAAGCAGCCCACTGATGCGCAAACTCAATACCAAGCCCAAGC 780
Db 6660 GAGATCAAACTCACAAAGCAGCCCACTGATGCGCAAACTCAATACCAAGCCCAAGC 6719
QY 781 AGTATGATGAGGACCTCGCAACATCCGCTCAGGGTCCGGAGAACGAGAAACCCACACA 840
Db 6720 AATGATGATGAGGACCTCATAACTCCGCTCAGGGTCCGGAGAACGAGAAACCCCTATACA 6779
QY 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACC 900
Db 6780 ACTTCTGATGCGGTCACTAAGCAGGGCTTTTCATCAACAATGCCACCACTCCCTCACC 6839
QY 901 CAACCAAGCAGCCACAGAGGAGAAACACACAAACCACTTCCCAAGATGCTGTGACT 960
Db 6840 CAACCAAGCAGCCACAGAGGAGAAACACACAGACCATTCCCAAGGTAAGTGTGACT 6899
QY 961 GAACTAGACAAATAAACAACATGTCACAAACCGTCCATGCCCTCTCATACACTACCA 1020
Db 6900 GAACTAGACAAATAAACAACATGTCACAAACCGTCCATGCCCTCTCATACACTACCA 6959
QY 1021 ATCTCTACTAACAACACTCCAAACACAACTTCAGCACTCTCTGACCACTTACAAAC 1080
Db 6960 ATCTCTACTAACAACACTCCAAAGCAAACTTCAGCACTCTCTGATACACTACAAAC 7019
QY 1081 ACCACAAATGACACACAGAGCAAACTCACTGAAATGAGCAACCACTGCGCCCTCG 1140
Db 7020 ACCACAAATGACACACAGAGCAAACTCACTGAAATGAGCAACCACTGCGCCCTCG 7079
QY 1141 ATAAACACCTGCTCCAAACGGAATCCCAACAGCAAGAGCAACCAAGCAAGCAAA 1200
Db 7080 AAAACACCTGCTCCAAACGGAATCTTACCACAGCAAGAGCACTTAAACACAGAA 7139
QY 1201 GGGCCCGGACAAACGGGACCAACAGCAAAATGAGCAATTTTCCAGCTCTCCCGCCACC 1260
Db 7140 GGGCCCGGACAAACGGGACCAAAATGAGCAAAATGAGCAATTTTAAACAGTCCCTCCCGCCACC 7199
QY 1261 CCCAGCTCGACTGACAAACATCTGTATATTTTCAAGAAAGCGATCGATTTCTCGGAA 1320
Db 7200 CCCAACCAGCACAAACATCTGTATATTTTCAAGAAAGCAAGTATCTCTCGGAGG 7259
QY 1321 GAAAGTGATATATCCCGTTTTTGTAGTGGTTTAAATAATGATGAAATGATTTTGCATCCA 1380
Db 7260 GAAAGTGATATATCCCGTTTTTGTAGTGGTTTAAATAATGATGAAATGATTTGATCCA 7319
QY 1381 ATCCCAACACAGAAACAACTTTGATGAATCTCCAGCTTTAATPACTTCAACTAATGAG 1440
Db 7320 GTTCCAAATACAAAGACGATCTTTGATGAATCTTCTAGTTCTGTGCTTCGGCTGAGAA 7379
QY 1441 GAAACACACCTCCCGGAAATCAGTTTAACTTTCTTATTTTCTGATTAATAATGGA 1500
Db 7380 GATCAACATGCTCCCGGAAATCAGTTTAACTTTCTTATTTTCTGATTAATAATGGA 7439
QY 1501 GATCTGCTACTCTGGGGAACACAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 7440 RACATGCTACTCTGGGGAACATGAAACGATTTGATGATGATGATGATGATGATGATGATGAT 7499
QY 1561 GTGACAGGAGGACGATTTGGCGGAGGGCTTAGCTGATACCAATTTTTTTGGCCCTCGAATC 1620

Db 7500 GTTCAGGAGGATGACCTGGCAGAGGGCTCAGTTGGATACCGCTTTTTTGGCCCTCGAATC 7559
QY 1621 GAAGGACTCTATCTATCGCGGTTTAAATCAAAAATCAGAAACAATTTAGTTTGTAGGTTGAGG 1680
Db 7560 GAAGGACTTTATCTGCTGGTTTAAATTAATAAACCAAAACAATTTGGTCTGCAGGTTGAGG 7619
QY 1681 CGTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGTCACAAACGAGAA 1740
Db 7620 CGTCTAGCCAAATCAAACTGCGCAATCTTGGAACTCTTATTAAAGAGTCACAAACGAGAA 7679
QY 1741 AGGACATTTCTTAACTCAATAGCAGCATCTGACTTTTGGCTTACGAGGTGGGCGGA 1800
Db 7680 AGGACATTTCTTAACTCAATAGCAGCATCTGACTTTTGGCTTACGAGGTGGGCGGA 7739
QY 1801 ACATGCAAGGTGCTAGGACCTGATTGTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
Db 7740 ACATGCAAGGTGCTAGGACCTGATTGTTGTCATAGGAATAGAGATCTTCCAGGAATAT 7799
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAGAGGAACTGGCTGGGCTCTA 1920
Db 7800 TCGGAACAAATTCACCAATCAAAAAGATGAAACAAAGAGGAGCTGGTTGGGCTCTA 7859
QY 1921 GGTGCAAAATGGTGGACATCTGACTGGGCTGTTCTCACCAATTTGGGCTCTGCTACTA 1980
Db 7860 GGTGTAATGGTGGACATCTGACTGGGCTGTTCTTACTAACTTGGGCTATTTGCTACTA 7919
QY 1981 TTATCTATAGCTGTTCTGATTGCTCTGCTCTGTATCTCTGCTATCTTCACTAAATACAT 2040
Db 7920 TTATCCATAGCTGTTGATTGCTCTATCTGATTTGCTGATTTGCTGATTTTACCAATATATC 7979
QY 2041 GGATGA 2046
Db 7980 GGGTAA 7985

RESULT 11
ABT13456
ID ABT13456 standard; DNA; 2046 BP.
XX AC ABT13456;
XX AC ABT13456;
DT 30-JAN-2003 (first entry)
XX DNA encoding a chimeric filovirus protein RVN-GP1/GP2.
DE Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX Unidentified.
OS Chimeric.
XX WO200279239-A2.
XX 10-OCT-2002.
XX 31-JAN-2002; 2002WO-US003339.
XX 31-JAN-2001; 2001US-0267522P.
PR (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
PA Grogan CC, Hevey MC, Schmaljohn AL;
PI WPI; 2003-040651/03.
DR P-FSDB; ABU18478.
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
XX Disclosure; Page 89-91; 94pp; English.
PS

CC The invention relates to a chimeric filovirus glycoprotein (GP) protein
CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
CC of GP2. The chimeric filovirus GP protein is useful for inducing an
CC immune response against infection of different filoviruses, specifically
CC against both Ebola and Marburg viruses by being used as a vaccine. This
CC polynucleotide sequence represents a DNA encoding a chimeric filovirus
CC protein of the invention
XX

SQ Sequence 2046 BP; 668 A; 481 C; 383 G; 514 T; 0 U; 0 Other;

Query Match	74.2%;	Score 1518;	DB 10;	Length 2046;
Best Local Similarity	83.9%;	Pred. No. 0;	Mismatches 330;	Indels 0; Gaps 0;
Matches 1716;	Conservative 0;			
QY 1 ATGAAGACACATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCC 60				
DB 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAACTCTCC 60				
QY 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGTCGGAACT 120				
DB 61 GTTTTAGAAATGCTAGTAACAGCCAACTCAAGATGTAGATTGATGTCCTCCGGAACC 120				
QY 121 CTCAGAGACAGACAGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAATTCCT 180				
DB 121 CTCCAAAAGACAGAGATGTTTCATCTGATGGGATTTACATGAGTGGGCAAAAATTCCT 180				
QY 181 GATTCCTTTTGGAGCATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240				
DB 181 GATTCCTTTTGGAGCATCTAAAGTGGGCTTTTCAGGACAGGTGTTCCTCCCAAGAA 240				
QY 241 GTTGTAGTACAGAGGGGAGGAGGCAAAACATGCTACAATATAGTGTAAACGGATCCC 300				
DB 241 GTTGTAGTACAGGAGGAGGAGGCAAAACATGTTTACAATATAGTGTAAACGACCT 300				
QY 301 TCTGGAATTCCTTGTGTTAGATCCTCTACCAATCCGTGACTATCCGNAATGCAAA 360				
DB 301 TCTGGAATTCCTTGTGTTAGATCCTCTACCAATCCGTGACTATCCGNAATGTA 360				
QY 361 ACTATCCATATATCAAGGTCAAAACCTCATGACAGGGGATCGCCTTCATTTATGG 420				
DB 361 ACTGTTTCAATATTCAGGTCAAAACCTCATGACAGGGGATTCGCTTCATTTATGG 420				
QY 421 GGAGCATTTTTCTGTATGATCGCAATGCTCCCAACAATGTAACGAGGCAAGTCTTC 480				
DB 421 GGAGCATTTTTCTGTATGATCGCTCTCAACAATGTAACGAGGCAAGTCTTC 480				
QY 481 ACTGAAGGAAATATAGCAGCTATGATTTGTAATAGACAGATTCACAGAAATGATTTTCT 540				
DB 481 ACTGAAGGAAATATAGCAGCTATGATTTGTAATAGACAGATTCACAGAAATGATTTTCT 540				
QY 541 CGGCAAGGACAAAGGTACCGTCTATATGAAATCTGACTTCTACTAATAATATTGGACAAGT 600				
DB 541 AGGCAAGGACAAAGGTATCTGATCAATGAATTTGACCTCCACCAATATATTGGACAAG 600				
QY 601 AGTAACGGAACGCAACGAATGACACTGGATGTTTGGCGCTCTTCAAGAAATACAATTC 660				
DB 601 AGCAATGAACGCAAGAAATGATCGGATGTTTGGCATCTCCAGAAATACAATTC 660				
QY 661 ACAAGAACCAACATGCTGCTGGTCCAAATATCCTCCACCACTGCCCCACAGCCGTCGG 720				
DB 661 ACAACCAATCAACACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAATCCG 720				
QY 721 GAGATCAAACTCACAGACCCCACTGATGACCCCAAACTCAATACACGAGCCCAAGC 780				
DB 721 AGCATTCATCTACAATATCTCAATTAATATCTGCTAAATCTGGAACTATGAACCAAGT 780				
QY 781 AGTGATGATGAGGACCTCGCAATCCGCTCAGGGTCGAGGAGCGAGAACCCCAACACA 840				
DB 781 AGCGAGATGAGGACCTTATGATTTCCGCTCAGGATCTGGAGACAGGGGCCCCACACA 840				
QY 841 ACTTCTGATGCGGTACACAGCAGGCGCTTTCATCAAGAAATGCCACCACTCCCTCACCA 900				
DB 841 ACTCTTAATGTAGTCACTGAACAGAAACAACTCGTCAACCAATATGTGTCCACTCTCTCACTA 900				

QY 901 CAACCAAGCAGCCGACAGCAAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 960				
DB 901 CATCCAAGCACCTCACAACATGAGCAAAACAGTAGCAATCTTCCCGACATGCTGTAACT 960				
QY 961 GAACTAGACAAAATAACAACTGCAACACGGTCCATGCCCCCTCAACACTACCACA 1020				
DB 961 GAGCACAATGGAAACCGACCCCAACCAACCAACAGCAAGCTCTCTCAACAACTAATACA 1020				
QY 1021 ATCTCTACTAAACACACCTCCAAACACAACTTCAGCACTCTCTGTGACCAATTAACAAAC 1080				
DB 1021 ACTCCACCTTAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCCAACCGCAAC 1080				
QY 1081 ACCCAATATGACACACAGAGCACAACTCACTGAAATAGGCAAAACAGTGCCCCCTCG 1140				
DB 1081 ATCAACCAATATGATACACACGCTGAACCTAGCAGAAAGCAACCAACCAATGCTCAGTTG 1140				
QY 1141 ATAACAACCTGCTCCACCGGAAATCCCAACCAACGCAAGAGACAGCAGCAAAA 1200				
DB 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACAGACAGACACCAACAGCAAC 1200				
QY 1201 GGCCCGCACACACGGCACCAACACAGCAAAATGAGCAATTTCCACAGTCTCTCCCCCACC 1260				
DB 1201 AACATCATGACGACATCAGATATACAGCAACACCCCAACAAATCTTCTCCGAT 1260				
QY 1261 CCCAGCTCGACTGCAACACATCTTGTATATTTCAAGAAAGCGATCGATTTCTGGAAA 1320				
DB 1261 TCTAGTCGCAACACCGGCCCTCTATATATCTTTAGAAAGAAACGATCGATTTCTGGAAA 1320				
QY 1321 GAAGGTGATATATTTCCCGTTTTTAGATGGGTTAATAACTGTAATTTGATTTGATCCA 1380				
DB 1321 GAAGGTGATATATTTCCCGTTTTTAGATGGGTTAATAACTGTAATTTGATTTGATCCA 1380				
QY 1381 ATCCCAACACAGAAACAACTCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1440				
DB 1381 ATCCCAACACAGAAACAACTCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1440				
QY 1441 GAAACAACACTCCCGCGAATATCAGTTTAACTTTCTTATTTTCTGATAAAATGGA 1500				
DB 1441 GAAACAACACTCCCGCGAATATCAGTTTAACTTTCTTATTTTCTGATAAAATGGA 1500				
QY 1501 GATCTGCTACTCTGGGAAAAACGAGAAATGATGTGATGACAGATTGAGGATTTGGAGT 1560				
DB 1501 GATCTGCTACTCTGGGAAAAACGAGAAATGATGTGATGACAGATTGAGGATTTGGAGT 1560				
QY 1561 GTGAGGAGGACGATTTGGCGGACGGCTTAGTGATACCAATTTTGGCCCTCGAATC 1620				
DB 1561 GTGAGGAGGACGATTTGGCGGACGGCTTAGTGATACCAATTTTGGCCCTCGAATC 1620				
QY 1621 GAAGGACTCTATCTGCGGTTTTAATCAAAATCAGAAACAATTTAGTTTGTAGTTGAGG 1680				
DB 1621 GAAGGACTCTATCTGCGGTTTTAATCAAAATCAGAAACAATTTAGTTTGTAGTTGAGG 1680				
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTGTTAAGGTCACACCCGAGAA 1740				
DB 1681 CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTGTTAAGGTCACACCCGAGAA 1740				
QY 1741 AGGACATTTCTTAATCAATAGGCATCAATTTGATCTTTTGTCTTACGAGGTGGGCGGA 1800				
DB 1741 AGGACATTTCTTAATCAATAGGCATCAATTTGATCTTTTGTCTTACGAGGTGGGCGGA 1800				
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860				
DB 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860				
QY 1861 TCAGAAACAAATCGACAAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGCTTA 1920				
DB 1861 TCAGAAACAAATCGACAAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGCTTA 1920				
QY 1921 GGTGCAAAATGCTGACATCTGACCTGGGTGTTCTACCAATTTGGGATCTCTGTACTTA 1980				
DB 1921 GGTGCAAAATGCTGACATCTGACCTGGGTGTTCTACCAATTTGGGATCTCTGTACTTA 1980				

Qy	1981	TTATCTATAGCTGTTCTGATTCCTCTCTGTCCTGTATCTGTCGTATCTTCACTAATAACATT	2040
Db	1981	TTATCTATAGCTGTTCTGATTCCTCTCTGTCCTGTATCTGTCGTATCTTCACTAATAACATT	2040
Qy	2041	GGATGA 2046	
Db	2041	GGATGA 2046	
 RESULT 12			
ID	AAD04043		
ID	AAD04043	standard; DNA; 2247 BP.	
XX	AAD04043;		
AC			
DT	02-JUL-2001	(first entry)	
XX			
DE	Marburg virus Ravn strain glycoprotein (GP) DNA.		
XX	Glycoprotein; GP; immune response; vaccine; antiviral;		
KW	type I transmembrane protein; ds.		
KW			
XX	Marburg virus.		
OS			
XX			
FH	Key	Location/Qualifiers	
FT	CDS	97..2142	
FT		/*tag= a	
FT		/product= "Marburg virus Ravn strain glycoprotein (GP)"	
XX			
PN	US6200959-B1.		
XX			
PD	13-MAR-2001.		
XX			
PF	04-DEC-1996;	96US-00760615.	
XX			
PR	04-DEC-1996;	96US-00760615.	
XX			
PA	(POWD-) POWDERJECT VACCINES INC.		
XX			
PI	Haynes JR, Schmaljohn CS, Fuller DL, Schmaljohn A, Jahrling PB;		
XX			
DR	WPI; 2001-280564/29.		
DR	P-PSDB; AAE00708.		
XX			
PT	Inducing an immune response to a Marburg or Ebola virus involves		
PT	delivering a genetic vaccine, which contains a genetic construct encoding		
PT	antigenic determinants for filovirus, using a particle acceleration		
PT	device.		
XX			
XX	Claim 6; Col 33-40; 33pp; English.		
CC	The patent discloses a method of inducing an immune response to Marburg		
CC	or Ebola virus glycoprotein which involves delivering a genetic vaccine,		
CC	containing a genetic construct encoding antigenic determinants for		
CC	filovirus, using a particle acceleration device. The genetic vaccine for		
CC	filovirus is created by joining a DNA sequence encoding at least a		
CC	portion of the filovirus glycoprotein to a promoter effective to promote		
CC	transcription of the DNA sequence. This method is useful for inducing		
CC	humoral, cell-mediated and secretory immune responses in the treated		
CC	individual. The present sequence is a DNA encoding Marburg virus Ravn		
CC	strain glycoprotein (GP), which is a type I transmembrane protein. This		
CC	sequence is used in the construction of genetic vaccine against		
CC	filoviruses		
XX			
SQ	Sequence 2247 BP; 734 A; 518 C; 424 G; 571 T; 0 U; 0 Other;		
 Query Match 73.8%; Score 1510; DB 4; Length 2247;			
Best Local Similarity 83.6%; Pred. No. 0;			
Matches 1711; Conservative 0; Mismatches 335; Indels 0; Gaps 0;			
Qy	1	ATGAAGACCATGTTTCCTTATCAGTCTTATCTTAATTCACAGGACAAAAAATCTCCCC	60
Db	97	ATGAAGACCATATATTTCTGATTAGTCTCATTTTATCCAAAGTATAAAACTCTCCCT	156

```
QY 1141 ATAAACAACCTCGCTCCACGCGGAATCCACACAGCAAGAACGACCAAGCAGCAAAAA 1200
Db 1237 AACACAACTCTAGATCCACACAGAAAATCCACACAGGACAGACCAACACAGCAACC 1296
QY 1201 GGGCCCGCCACAAACGCGCACAAACAGCAAAATAGCATTTCCACAGTCCTCCGCCACC 1260
Db 1297 ACATCATCATCGACATCAGATATACAGCAACACCCACAAATCTCTCCGGAT 1356
QY 1261 CCAGCTCGACTGCACAAACATCTTGATATATTTAGAGAAAGCGATCGATTTCTGAAA 1320
Db 1357 TCTAGTCCGACAAACCGCCCTCCTATATCTTTAGAAAGAAACGAAGCATTTCTCGAAA 1416
QY 1321 GAAGTGATATATTTCCGTTTTAGATGGTTAAATAATAGTGAATTTGATGCA 1380
Db 1417 GAAGTGATATATTTCCGTTTTAGATGGTTAAATAATAGTGAATTTGATGCA 1476
QY 1381 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTAATACCTCAACTAATGAG 1440
Db 1477 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTAATACCTCAACTAATGAG 1536
QY 1441 GAACAACACATCCCGCGAATATCAGTTTAACTTTCTTATTTTCTGATAAAAAATGGA 1500
Db 1537 GAACAACACATCCCGCGAATATCAGTTTAACTTTCTTATTTTCTGATAAAAAATGGA 1596
QY 1501 GATAGCTCTACTCTGGGAAAAACGAGATGATTTGTGATGAGATGAGGATTTGGAGT 1560
Db 1597 GATAGCTCTACTCTGGGAAAAACGAGATGATTTGTGATGAGATGAGGATTTGGAGT 1656
QY 1561 GTGAGGAGGACGATTTGGCGCAGGCTTACGTGATACCAATTTTGGCCCTCGAATC 1620
Db 1657 GTGAGGAGGACGATTTGGCGCAGGCTTACGTGATACCAATTTTGGCCCTCGAATC 1716
QY 1621 GAAGGACTCTATCTGCGGTTTAAATCAAAAAATCAGAAATATAGTTTGTAGTTGAGG 1680
Db 1717 GAAGGACTCTATCTGCGGTTTAAATCAAAAAATCAGAAATATAGTTTGTAGTTGAGG 1776
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCTCTGGAGCTCTGTTAAGGTCACAAACGAGAA 1740
Db 1777 CGCTTAGCTAATCAAACTGCTAAATCTCTGGAGCTCTGTTAAGGTCACAAACGAGAA 1836
QY 1741 AGGACATTTTCTTAAATCAATAGGACATGCAATTTGCTTACGAGTGGGGCGGA 1800
Db 1837 AGGACATTTTCTTAAATCAATAGGACATGCAATTTGCTTACGAGTGGGGCGGA 1896
QY 1801 ACATGCAAGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
Db 1897 ACATGCAAGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1956
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAAGGAGGAACTGGCTGGGGTCTA 1920
Db 1957 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAAGGAGGAACTGGCTGGGGTCTA 2016
QY 1921 GGTGGCAATGGTGACATCTGACTGGGTGTTCTCACCATAATTCGGCATCTGCTACTA 1980
Db 2017 GGTGGCAATGGTGACATCTGACTGGGTGTTCTCACCATAATTCGGCATCTGCTACTA 2076
QY 1981 TTATCTATAGCTGTTCTGATGCTCTGCTCTGATCTGCTGATCTTCACTAAATACATT 2040
Db 2077 TTATCTATAGCTGTTCTGATGCTCTGCTCTGATCTGCTGATCTTCACTAAATACATT 2136
QY 2041 GGATGA 2046
Db 2137 GGATGA 2142
```

RESULT 13

ABT13453

ID ABT13453 standard; DNA; 2046 BP.

XX

AC ABT13453;

XX

DT 30-JAN-2003 (first entry)

XX

DE DNA encoding a chimeric filovirus protein RVN-GP1/MUS-GP2.

XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;

KW immune response; Ebola; Marburg virus; vaccine; gene; ds.

XX Unidentified.

OS Chimeric.

XX WO200279239-A2.

PN 10-OCT-2002.

XX 31-JAN-2002; 2002WO-US003339.

XX 31-JAN-2001; 2001US-0267522P.

XX (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.

PA Grogan CC, Hevey MC, Schmaljohn AL;

XX WPI; 2003-040651/03.

XX P-PSDB; ABJ18475.

XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,

PT useful for inducing an immune response against infection of different

PT filoviruses, specifically against both Ebola and Marburg viruses.

XX Claim 18; Page 78-80; 94pp; English.

XX The invention relates to a chimeric filovirus glycoprotein (GP) protein

CC comprising GP1 and GP2, where GP1 is from a filovirus different than that

CC of GP2. The chimeric filovirus GP protein is useful for inducing an

CC immune response against infection of different filoviruses, specifically

CC against both Ebola and Marburg viruses by being used as a vaccine. This

CC polynucleotide sequence represents a DNA encoding a chimeric filovirus

CC protein of the invention

XX Sequence 2046 BP; 662 A; 486 C; 379 G; 519 T; 0 U; 0 Other;

SQ Query Match 64.5%; Score 1319.6; DB 10; Length 2046;

Beat Local Similarity 77.8%; Pred. No. 0;

Matches 1592; Conservative 0; Mismatches 454; Indels 0; Gaps 0;

QY 1 ATGAAGACCACATGTTCTTATCAGTCTTATCTTAATCAAGGACAAAAATCTCCC 60

Db 1 ATGAAGACCACATGTTCTTATCAGTCTTATCTTAATCAAGGACAAAAATCTCCC 60

QY 61 ATTTTAGAGTAGCTAGTAAATCAACCCCAAAATGTGGATTGCTCGGAACT 120

Db 61 GTTTTAGAAATTTGCTAGTAAACAGCCAACTCAAGATGTAGATTGCTCGGAACT 120

QY 121 CTCAGAAAGACAGAACGCTCCATCTGATGGGATTCACACTGAGTGGCAAAAAGTTGCT 180

Db 121 CTCAGAAAGACAGAACGCTCCATCTGATGGGATTCACACTGAGTGGCAAAAAGTTGCT 180

QY 181 GATTCCCTTTGGAGGCATCAAGCATGGGCTTTACAGACAGGTGTACTCCCAAGAAAT 240

Db 181 GATTCCCTTTGGAGGCATCAAGCATGGGCTTTACAGACAGGTGTACTCCCAAGAAAT 240

QY 241 GTTCAGTACACAGGGGGAGAGCCAAACATGCTACATATAGTGAACGATCC 300

Db 241 GTTCAGTACACAGGGGGAGAGCCAAACATGCTACATATAGTGAACGATCC 300

QY 301 TCTGAAAAATCTTGTGTTAGATCTCTTACCAACATCCCGTACTATCCGAAATGCAAA 360

Db 301 TCTGAAAAATCTTGTGTTAGATCTCTTACCAACATCCCGTACTATCCGAAATGCAAA 360

QY 361 ACTATCCATCATATTCAGGTCAAAAACCTCATGACAGGGGATCGCCCTTCATTATGG 420

Db 361 ACTATCCATCATATTCAGGTCAAAAACCTCATGACAGGGGATCGCCCTTCATTATGG 420

QY 421 GGAGCATTTTCTGTATGATCGCATTCCTCCCAACAATGATCCGAGGAAAGTCTTC 480

Db 421 GGGGCAATTTTCTTGATGATCGGTTGCTCTACAACAATGTACCGAGCAAGTCTTC 480
Qy 481 ACTGAAGGAAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAAATGATTTCTCG 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATTTGTAAATAGACAGTTCACAGAATGATTTTCT 540
Qy 541 CGGCAAGCACAGGTCACCGTCATATGAATCTGACTTCTACTAATAATATTTGACAGT 600
Db 541 AGGCAGGACAGGTTATCGTCACATGAACCTTGACTCCACCAATAAATATTTGACAGC 600
Qy 601 AGTAAACGGAACGCAACGAATGACACTGATGTTTTCGGCGCTCTTCAAGAAATACAATTC 660
Db 601 AGCAATGAAACGACAGAAATGATAGGGAATGTTTGGCATCTCCAAAGAAATACAATCC 660
Qy 661 ACAAGAACCAACATGTCCTCGTCCAAATACCTCCACACATGCCACAGCCGCTCG 720
Db 661 ACAACAATCAACATGCGCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAATCCG 720
Qy 721 GAGATCAAACTCACAGCACCCCAACTGATGCCAACCAACTCAATACCACGAGCCCAAGC 780
Db 721 AGCATTCACTCTACAATACTCAAAATTAATCTGCTAAATCTGGAACATATGAACCCAGT 780
Qy 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACAGAAACCCACACA 840
Db 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCAACA 840
Qy 841 ACTTCTGATGGGTACAACGAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA 900
Db 841 ACTCTTAATGTAGTCACTGACACAGAAACAAATCGTCAACAATATTTGCTCCACTCTCACTA 900
Qy 901 CAACCAAGCAGCCACAGCAGGAGGAAACACACACACACACACACACACACACACAC 960
Db 901 CATCCAAGCATCTCACACATGAGCAAAACAGTAGACGAATCTTCCCGACATGCTGTAACT 960
Qy 961 GAACTAGACAAAATAACACAACTGCAACACCGTCCATGCCCTCCATACCACTACCA 1020
Db 961 GAGCACATGNAACGACCCACACACACACACACACACACACACACACACACACACAC 1020
Qy 1021 ATCTCTACTAACACACCTCCAAACACAACTTTCAGCAGCTCTCTGTGACCAATTAACAAAC 1080
Db 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCCAACCGCAAC 1080
Qy 1081 ACCACCAATGACACACAGCAGCAGACATCACTGAAATGAGCAAAACGAGTCCCGCTCG 1140
Db 1081 ATCAACAATATGATACACACACGTAAGTACAGAAAGCGAACCAACCAATGCTCAGTTG 1140
Qy 1141 ATAACAACCTCGCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCACAGCAAAACAAA 1200
Db 1141 AACACAATCTAGATCCAAACAGAAATCCCAACACAGCAGCAAGACACCAACAGCACACC 1200
Qy 1201 GGCCCCGCCAACAGGCAACCAACACGACAAATGAGCATTTTCAACAGTCTCTCCCGCCACC 1260
Db 1201 AACATCATCATGACGACATCAGATATAACAAGCAACACCCCAACAAATTTCTCTCGGAT 1260
Qy 1261 CCCAGCTCGACTGACACACATCTGTATATTTTCAGAGAAAGCGATCGATTTCTGGAAC 1320
Db 1261 TCTAGTCCGCAACCGCCCTCTATATACTTATAGAAAGAAACGATCGATCTCTGGAGG 1320
Qy 1321 GAAGTGATATATTCCTGTTTACATGGGTTTAAATAATCTGAAATGATTTTGTATCCA 1380
Db 1321 GAAGCGACATGTTCCCTTTCTGATGGGTTTAAATGATCTCCAAATGATTTTGAACCCA 1380
Qy 1381 ATCCCAACACAGAAACAACTTTTGTATGAATCTCCAGCTTTAATFACTTCAACTAATGAG 1440
Db 1381 GTTCCAAATACAAAACAACTCTTGATGAATCTCTAGTTCTGTGTCCTCGCTGAGGAA 1440
Qy 1441 GAAACACACATCCCGGAAATATGATTTTCTTATTTTCTGATTAATAAATGGA 1500
Db 1441 GATCAACATGCTCCCGCAATATAGTTTAACTTTATCTTATTTCTTAATATAAATGAG 1500
Qy 1501 GATCTGCTACTCTGGGGAACCAAGAAATGATTTGTGATGACAGAGTTTCAGGATTTGGAGT 1560
Db 1501 AACCTGCTACTCTGGAGAAATGAGAAATGATTTGTGATGACAGAGTTAAGAAATTTGGAGC 1560

Qy 1561 GTGCAGAGAGACGATTTGGCGGAGGCTTAGCTGGATACCAATTTTGGCCCTGGAAATC 1620
Db 1561 GTTCAGAGAGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTGGCCCTGGAAT 1620
Qy 1621 GAAGGACTCTATCTGCGGGTTTAAATCAAAAATCAGAAACAATTTAGTTTGTAGGTTGAGG 1680
Db 1621 GAAGGACTTTTACTGCTGCTGTTTAAATTAATAAATCAAAAACAATTTGGTCTGCAGGTTGAGG 1680
Qy 1681 CGCTTACTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGTCACAAACCGAGAA 1740
Db 1681 CGCTAGCCCAATCAAACTGCAAAATCTTGGAACTCTTATTTAGAGATCACAACCTGAGGAA 1740
Qy 1741 AGGACATTTCTTAAATCAATAGCATGCTATTTGACTTTCTACTCACAAGATGGGGAGGA 1800
Db 1741 AGAATCTTCTTAAATCAATAGCATGCTATTTGACTTTCTACTCACAAGATGGGGAGGA 1800
Qy 1801 ACATGCAAGTGTCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAAATATC 1860
Db 1801 ACATGCAAGTGTCTAGGACCTGATTTGTCATGCGGATAGAGACTTGTTCNAAAATAT 1860
Qy 1861 TCAGAAACAAATCGACAAATCAGAAAGGATGAACAAAGGAGGAACTGGCTGGGCTA 1920
Db 1861 TCAGAGCAAAATGACCAAAATTAAGAGAGACGACAAAGAGGGGACTGGTTGGGCTG 1920
Qy 1921 GGTGGCAAAATGGTGACATCTGACTGGGGTGTCTTCAACAAATTTGGGCACTCTGCTACTA 1980
Db 1921 GGTGGTAAATGGTGACATCTGACTGGGGTGTCTTCTTAACCTTGGGCACTTTTCTACTA 1980
Qy 1981 TTATCTAGTGTCTGATTTGCTCTGCTGCTGATCTGCTATCTTCACTTAAATACATT 2040
Db 1981 TTATCCATAGTGTCTGATTTGCTCTATCTCTGATTTGCTATCTTTACTTAAATATATC 2040
Qy 2041 GGATGA 2046
Db 2041 GGATGA 2046
RESULT 14
ABT13451
ID ABT13451 standard; DNA; 1841 BP.
XX
AC ABT13451;
XX
DT 30-JAN-2003 (first entry)
XX
DE DNA encoding a chimeric filovirus protein MBGV-GP1/EBOV-GP2.
XX
KW Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX
OS Unidentified.
OS Chimeric.
XX
PN WO200279239-A2.
XX
PD 10-OCT-2002.
XX
PF 31-JAN-2002; 2002WO-US003339.
XX
PR 31-JAN-2001; 2001US-0267522P.
XX
PA (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
XX
PI Grogan CC, Hevey MC, Schmaljohn AL;
XX
XX WPI; 2003-040651/03.
DR P-ESDB; ABJ18473.
XX
PT New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
XX

PS Claim 14; Page 70-72; 94pp; English.

XX The invention relates to a chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2, where GP1 is from a filovirus different than that of GP2. The chimeric filovirus GP protein is useful for inducing an immune response against infection of different filoviruses, specifically against both Ebola and Marburg viruses by being used as a vaccine. This CC polynucleotide sequence represents a DNA encoding a chimeric filovirus protein of the invention

SQ Sequence 1841 BP; 589 A; 500 C; 361 G; 391 T; 0 U; 0 Other;

Query Match 63.9%; Score 1308; DB 10; Length 1841;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACACATGTTCCCTTATCAGTCTTATCTTAATTCAGGGACAAAATCTCCCC 60
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1 ATGAGACACATGTTCCCTTATCAGTCTTATCTTAATTCAGGGACAAAATCTCCCC 60
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 ATTTTAGATAGTAGTAATTAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAAT 120
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 ATTTTAGATAGTAGTAATTAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAAT 120
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 CTCGAGAGACAGAGACGTCCATCTGATGGATTCAACATGAGTGGGCAAAAGTTGCT 180
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 CTCGAGAGACAGAGACGTCCATCTGATGGATTCAACATGAGTGGGCAAAAGTTGCT 180
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 GATTCCTTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCGAAGAT 240
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 GATTCCTTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCGAAGAT 240
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 GTTGTAGTACAGAGGGGAGGAAGCCAAAACATGCTACAATATAAGTGAACGGATCCC 300
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 GTTGTAGTACAGAGGGGAGGAGGCCAAAACATGCTACATATAAGTGAACGGATCCC 300
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 TCTGAAAATCTCTGTGTAGATCTCTCAACAAATCCGTGACTATCCGAAATGCAAA 360
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 TCTGAAAATCTCTGTGTAGATCTCTCAACAAATCCGTGACTATCCGAAATGCAAA 360
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 ACTATCCATCATATTCMAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTATGG 420
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 ACTATCCATCATATTCMAGGTCAAAACCTCATGACAGGGGATGCCCTTCATTATGG 420
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 CGAGCATTTTCTGTATGATCGCTTGCTCCAAACAAATGCTACGAGGCAAAAGTCTTC 480
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 CGAGCATTTTCTGTATGATCGCTTGCTCCAAACAAATGCTACGAGGCAAAAGTCTTC 480
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 ACTGAAGGAAACATAGCAGCTATGATTTGTCATAAGACAGTGCACAAAATGATTTCTCG 540
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 ACTGAAGGAAACATAGCAGCTATGATTTGTCATAAGACAGTGCACAAAATGATTTCTCG 540
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 541 CGGCAAGGACAGGGTACCGTCATATGATCTGACTTCTACTAATAATATTTGGACAGT 600
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 541 CGGCAAGGACAGGGTACCGTCATATGATCTGACTTCTACTAATAATATTTGGACAGT 600
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 601 AGTAACGGAACGCAACAGTAACACTGATGATGTTTGGGGCTCTTCAAGATACAAATCT 660
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 601 AGTAACGGAACGCAACAGTAACACTGATGATGTTTGGGGCTCTTCAAGATACAAATCT 660
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 661 ACAAGAACCAACATGTCTCCGTCACAAAATACCTCCACCACTGCCACAGCCCGTCG 720
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 661 ACAAGAACCAACATGTCTCCGTCACAAAATACCTCCACCACTGCCACAGCCCGTCG 720
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 721 GAGATCAAACTACAAGACCCCACTGATGCGACCAAACTCAATACACGACCCCAAGC 780
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 721 GAGATCAAACTACAAGACCCCACTGATGCGACCAAACTCAATACACGACCCCAAGC 780
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 781 AGTGTATGATGAGGACCTCCGCTCAGGTCAGGGTCCGGAGACGAGAACCCACACACA 840
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 781 AGTGTATGATGAGGACCTCCGCTCAGGTCAGGGTCCGGAGACGAGAACCCACACACA 840
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAAATGCCACCCACTCCCTCACCA 900

DB 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAAATGCCACCCACTCCCTCACCA 900
QY 901 CAACCAAGCAGCCACAGCAGAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 960
DB 901 CAACCAAGCAGCCACAGCAGAGGAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 960
QY 961 GAACTAGACAAAATAAACAACAACTGCAACCGTCCATGCCCCCTCATAAACAACACTACCACA 1020
DB 961 GAACTAGACAAAATAAACAACAACTGCAACCGTCCATGCCCCCTCATAAACAACACTACCACA 1020
QY 1021 ATCTCTACTAACAACCTCCAAAACAACTTTCAGCACTCTCTCTGACCACTTACAAAAC 1080
DB 1021 ATCTCTACTAACAACCTCCAAAACAACTTTCAGCACTCTCTCTGACCACTTACAAAAC 1080
QY 1081 ACCACCAATGACACACAGAGCAGACATCACTGAAAATGAGCAAAACAGTGCCTCCCTCG 1140
DB 1081 ACCACCAATGACACACAGAGCAGACATCACTGAAAATGAGCAAAACAGTGCCTCCCTCG 1140
QY 1141 ATAAACAACCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAGCAGCAAGAAAAA 1200
DB 1141 ATAAACAACCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAGCAGCAAGAAAAA 1200
QY 1201 GGGCCCGCACAACGGGACCAACACGACCAATGAGCAATTCACCAAGTCTCTCCCGCACC 1260
DB 1201 GGGCCCGCACAACGGGACCAACACGACCAATGAGCAATTCACCAAGTCTCTCCCGCACC 1260
QY 1261 CCCAGCTCGACTGCACAACTCTTGTATATTTTCAGAAAGAAAGCGATCG 1308
DB 1261 CCCAGCTCGACTGCACAACTCTTGTATATTTTCAGAAAGAAAGCGATCG 1308

RESULT 15
ACC71562
ID ACC71562 standard; DNA; 6902 BP.
XX
AC ACC71562;
XX
DT 10-JUL-2003 (first entry)
XX
DE VRC6703(pVR1012x/s Marburgdelta TM/h (codon optimised)) plasmid.
XX
KW Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
OS Synthetic.
XX
PN WO2003028632-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030251.
XX
PR 01-OCT-2001; 2001US-0326476P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Nabel GJ, Yang Z, Sullivan N, Sanchez A;
PI WPI; 2003-371961/35.
DR
PT New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
PS Claim 1; Page 213-215; 219pp; English.
XX
CC The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprising of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza

CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
XX virus infection

SQ Sequence 6902 BP; 1753 A; 1978 C; 1707 G; 1464 T; 0 U; 0 Other;

Query Match 48.7%; Score 995.6; DB 10; Length 6902;

Best Local Similarity 69.5%; Pred. No. 2e-275;

Matches 1352; Conservative 0; Mismatches 594; Indels 0; Gaps 0;

Qy	1	ATGAAGACCATGTTCTCTATCATGCTTATCTTAATTCAGGGACAAAAATCTCCC	60
Db	1923	ATGAAGACCATGTTCTCTATCATGCTTATCTCTGATCCAGGGCATCAAGACCTGCCC	1982
Qy	61	ATTTTATGAGATGCTAGTAATTAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAAT	120
Db	1983	ATCTGGAGATCGCCAGCAACACAGCCCCAGAACGTGGACGGTGTGCAGCGGCACC	2042
Qy	121	CTCAGAGACAGAAAGACGTCCATCTGATGGGATTTCACACTGAGTGGGCAAAAATTGCT	180
Db	2043	CTGCAGAAAGACGAGGACGTGCACCTGATGGGCTTCACCTGAGCGGCAGAAAGTGGCC	2102
Qy	181	GATTCCTCTTGGGGCATCCAAGCGATGGGCTTTCAGGACAGGTGATCCTCCCAAGAT	240
Db	2103	GACAGCCCTCTGGAGGCCAGAAAGGTGGGCTTTCAGAGCGGGCTGCCCCCAAGAAC	2162
Qy	241	GTGTAGTACACAGAGGGGAGGAAGCCAAACATCTCAATATAGTGTAAACGATCCC	300
Db	2163	GTGGAGTACACCGAGGCGAGGAGCCAGACTCTGTACACATCAGGTGACGCCNCCC	2222
Qy	301	TCTGAAAAATCTTCTGTGTAGATCCTCTACCAACATCCGTGACTATCCGAATGCAAA	360
Db	2223	AGCGGCAAGAGCCTGTCTGTGACCTCCCAACACATCAGGACTACCTTAAGTGCAAG	2282
Qy	361	ACTATCCATATATTCAGGTCAAAACCTCATGCAAGGGGATCGCCCTTCATTTATGG	420
Db	2283	ACCATCCACCATCCAGGGCCAGAACCTTCAGGCCAGGGGATCGCCCTGCACCTGTGG	2342
Qy	421	GGAGCATTTTCTGTATGATCGCATGCTCCACAAATGTATCCGAGGCAAAAGTCTTC	480
Db	2343	GGCGCTTCTTCTGTACGACAGGATCGCAGCACACATGTACAGGGGAGGSGTTC	2402
Qy	481	ACTGAAGGGAACATAGCAGCTATGATGTCAATAGACAGTGCACAAAATGATTTCTCG	540
Db	2403	ACCGAGGGCAACATCGCGCCATGATCGTTAAACAGACCGTGCAACAGATGATCTCAGC	2462
Qy	541	CGGCAAGGACAGGATACCGTATATGATCTGACTTCTAATAAATATTTGGAAGT	600
Db	2463	AGGCAAGGCGCCAGGGGTACAGGCATGAACTTGACAGCACCAACAGTACTTGGACCCAGC	2522
Qy	601	AGTAAACGGAACCAACAGTAACACTGGATGTTTGGCGCTCTTCAAGAAATACAAATCT	660
Db	2523	AACAAGGCAACAGACCAACAGACCGGCTGCTTGGGGCCCTGACAGAGTACAAACAGC	2582
Qy	661	ACAAAGAACCAAAATGTGCTCCGTCCAAAATACCTTCCACCACTGCCACAGCCCGTCCG	720
Db	2583	ACCAAGAACCAACCTGCGCCCCCAGCAAGATCTCCAGCCCTGCCCCAGCCAGGGCCC	2642
Qy	721	GAGATCAAACTCACAGACCCCAACTGATGCCACCAACTCAATACACAGGACCCCAAGC	780
Db	2643	GAGATCAAGCCACCAAGACCCCCACCGACGCCACCACTTGAACACCAACGACCCCAAC	2702
Qy	781	AGTGTATGATGAGGACTTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA	840
Db	2703	AACGACGACGAGGACTGTATACACAGCGGACCGGACGGGCGAGGAGCCCTTACACC	2762

Qy	841	ACTTCTGATGGGTCCACCAAGCAGGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA	900
Db	2763	ACCAAGCAGCGCGTGACCAAGCGGGCTTGAGCAGCACCATTGCTCTTACCCCTAGCCCT	2822
Qy	901	CAACCAAGCAGCGCACAGCAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT	960
Db	2823	CAGCCAGCACCTCTAGCAGGAGGGCAACACACCGACCAAGCAGCAGGCGGACCGTGACC	2882
Qy	961	GAACTAGACAAAAATAACACAACTGTCACAAACCGTCCATGCCCCCTCATACACTACCA	1020
Db	2883	GAGCCCAACAGACCAACACACCGCCAGCCAGCATGCTCTCTCAACACACACCGCC	2942
Qy	1021	ATCTCTACTTAACAACCTCCAAACACAACTTCAGCACTCTCTCTGCACTTACAAAC	1080
Db	2943	ATCAGCACCAACCAACACAGCAAGAACAACTTCAGCACCTTGAGCGGTGAGCGCTGAG	3002
Qy	1081	ACCACCAATGACACACAGCAGCACAATCACTGAAATGAGCAACACAGTGCCTCCCTCG	1140
Db	3003	ACCACCAATGACACACAGCAGCACAACCGCACCGGAAACGAGCAGCAGCGCCCTAGC	3062
Qy	1141	ATAACAAACCTTCCCAACCGGAAATCCCAACAGCAAAAGAGCAGCAGCAAAAAA	1200
Db	3063	AAGACCACTTCCCTCCACCGCAACCTGACCAACCGCAAGAGCACCACCAACACCAAG	3122
Qy	1201	GGCCCCGACCAACGGGACCAAAACAGCAAAATGAGCAATTCACCACTCTCCCCCACC	1260
Db	3123	GGCCCCACACACCGCCCCCTAACTGACCAACCGCCACTGACAGGCCCCAGGCCACC	3182
Qy	1261	CCCAGCTGAGTGCACAACTCTTGTATATTTTCAAGAAAGAGCGATCGATTTCTGGAAA	1320
Db	3183	CCCAACCCCAACCCAGCACCTGTGTATCTTCAAGAAAGAGGAGGATCTGTGGAGG	3242
Qy	1321	GAAGGTGATATATTTCCGTTTTTATAGTGGTTTAAATACTGAAATTTGATTTTGATCCA	1380
Db	3243	GAGGCGGATATGTTTCCCTTCTGAGCGGCTGTATCAAGCCCTTATCGACTTCGACCCC	3302
Qy	1381	ATCCAAACACAGAAAACATCTTTTGTATGATCTCCAGCTTTAATACTCTCAACTAATGAG	1440
Db	3303	GTGCCCCAACCAAGACCATCTTCGACGAGACGACGACGCGCGCCGCGCAGGAG	3362
Qy	1441	GAAACACACATCCCGGAAATATCAGTTTAACTTTCTCTTATTTTCTGATAAAATGGA	1500
Db	3363	GACAGACAGCGCAGCCCAACATCAGCTGACCTGACCTTACTTCCCCCAACATCAACGAG	3422
Qy	1501	GATCTGCTACTCTCTGGGAAAAACAGAAATGATTTGTATGACAGATTTGAGGATTTGGAGT	1560
Db	3423	AACACCGCTTACAGCGCGAGACGAGAACGACTGCGACGCGGAGCTGAGGATCTGGAGC	3482
Qy	1561	GTGAGGAGGACGATTTGGCGGAGGGCTTATGCTGGATACCAATTTTGGCCCTCGAATC	1620
Db	3483	GTGAGGAGGACGACTGGCGCGGCGCTGAGCTGGATTTCCCTTCTCGGCCCGGCAATC	3542
Qy	1621	GAAGGACTATATCTGCGGTTTTTAAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG	1680
Db	3543	GAGGCGCTGTACCGCGGCTGTATCAAGAACCAAGAACACCTGGTGTGCAAGGCTGAGG	3602
Qy	1681	CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGGTCAACACCGAGAA	1740
Db	3603	AGGCTGGCCCAACAGACCGCCAGAGCCCTGGAGCTGCTGAGGGTGCACACCGAGGAG	3662
Qy	1741	AGGACATTTTCTTAATCAATAGGATGCAATTTGATTTTTTGTCTTACGAGGTGGGGCGA	1800
Db	3663	AGGACCTTTCAGCTGTATCAACAGGACCGCATCGACTTCTCTGTCAGCAGGTGGGGCGC	3722
Qy	1801	ACATGCAAGGTGCTAGGACCTGATTTGTGCTAGGAATAGAGATCTATCTATAAAATATC	1860
Db	3723	ACCTGCAAGGTGCTGGGCCCGACTGTGCTATCGGATCGAGGACTGAGCGAGAACATC	3782
Qy	1861	TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAAGGAGGAACTGGCTGGGGTCTA	1920
Db	3783	AGCGAGCAGATCGACCATCAAGAGGACGAGCAGAGGAGGACCGGCTGGGGCGCTG	3842
Qy	1921	GGTGCAAAATGGTGACATCTGACTG	1946

Db 3843 GCGGCAAGTGTGGACCAGCGACTG 3868
|| ||||| ||||| |||||

Search completed: July 31, 2005, 16:57:23
Job time : 1078.87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:13:05 ; Search time 6679.18 Seconds
(without alignments)
11660.032 Million cell updates/sec

Title: US-10-066-506A-5

Perfect score: 2046

Sequence: 1 atgaagaccacatgttcct.....tcactaaatacattggtgatga 2046

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	99	4.8	895	9	CNS0071A
C 2	93.2	4.6	1101	9	CNS00FXE
C 3	87.2	4.3	922	9	CNS0073W
C 4	83.4	4.1	884	9	CNS006U0
C 5	80.8	3.9	1101	9	CNS00LO0
C 6	80.6	3.9	1101	9	CNS00LT2
C 7	79.2	3.9	909	9	CNS00JTL
C 8	77.8	3.8	1225	9	CNS0166K
C 9	75	3.7	837	9	AG136151
C 10	73.6	3.6	1965	9	CL090566
C 11	72.6	3.5	939	9	CNS006U0
C 12	72.2	3.5	1787	9	CG754239
C 13	71.8	3.5	1626	9	AG131986
C 14	71.4	3.5	1455	9	AG382036
C 15	70	3.4	1101	9	CNS017YH
C 16	69.2	3.4	902	9	CNS006GP
C 17	68.4	3.3	1101	9	CNS017ZT
C 18	68.2	3.3	861	9	CNS0075A
C 19	68	3.3	939	9	CNS00CNG
C 20	67.4	3.3	1559	9	CG756460
C 21	67	3.3	796	8	B12592
C 22	67	3.3	993	9	CNS005N6
C 23	67	3.3	1319	9	AG371230
C 24	67	3.3	1667	9	CG756610

C	25	66.2	3.2	470	4	BG786319	SEUMC006
	26	66	3.2	723	8	AZ627893	1M0469118
	27	66	3.2	1025	9	CL487193	SAIL_449
	28	65.6	3.2	527	9	CE480293	tiGr-ges-
	29	65.6	3.2	860	9	CNS018FL	Drosophil
	30	65.4	3.2	913	6	CA986290	AGENCOURT
	31	65.4	3.2	925	7	CK425533	AUF IPtes
	32	65	3.2	769	8	BZ579780	mbh2_769
	33	65	3.2	869	9	AG137111	Pan trogl
	34	65	3.2	1147	8	CC187235	CH261-152
	35	64.8	3.2	513	9	CNS02CMF	AL191328
	36	64.8	3.2	633	7	CV070818	Tetraodon
	37	64.6	3.2	937	9	CNS006ST	CS_gil_34
	38	64.6	3.2	1153	9	CNS07BVP	AL065880
	39	64.6	3.2	1345	8	AQ743328	Drosophil
	40	64.6	3.2	1613	8	BZ557155	T3 end of
	41	64.4	3.1	917	9	CL466449	HS_5387_B
	42	64.2	3.1	1131	9	CNS03AWV	pacsl-60
	43	64	3.1	2263	3	CR698468	SAIL_1256
	44	63.8	3.1	732	9	AG135778	AL235768
	45	63.8	3.1	1183	9	AG136828	Tetraodon

ALIGNMENTS

RESULT 1
CNS0071A/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit fly); genomic survey sequence.

ACCESSION
AL066286
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

CNS0071A 895 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit fly); genomic survey sequence.
AL066286
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 895)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoler in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source

1..895
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14B09"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 4.8%; Score 99; DB 9; Length 895;
Best Local Similarity 24.1%; Pred. No. 2.8e-16;

[illegible]

RESULT 2	CNS00FXE	1101 bp	DNA	linear	GSS 03-JUN-1999
CNS00FXE/c					
LOCUS					
DEFINITION	Drosophila melanogaster genome survey sequence TtT3 end of BAC: BACR32C19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				

AL071370	GI:4951210				
AL071370.1					
GSS.					
Drosophila melanogaster (fruit fly)					
Drosophila melanogaster					
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;					
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;					
Ephydroidea; Drosophilidae; Drosophila.					
1 (bases 1 to 1101)					
Genoscope.					
Direct Submission					
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;					
BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr					
- Web : www.genoscope.cns.fr)					
Determination of this BAC-end sequence was carried out as part of a					
collaboration with the Berkeley Drosophila Genome Project (BDGP) .					
The BDGP is constructing a physical map of the Drosophila					
melanogaster genome using these BACs. For further information					
please see http://www.fruitfly.org The BDGP Drosophila					
melanogaster BAC library was prepared by Kazutoyo Osoegawa and					
Aaron Mammoser in Pieter de Jong's laboratory in the Department of					
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,					
NY. The library is named RPCT-98 and was constructed by partial					
EcoRI digestion of Drosophila DNA provided by the BDGP from the					
isogenic strain y2; cn bw sp, the same strain used for the BDGP's					
P1 and EST libraries. A more detailed description of the library					
and how to order individual BAC clones, the entire library, or					
filters for hybridization from the BACPAC Resource Center can be					
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .					

FEATURES	source
found at	location/Qualifiers
	1. 1101
	/organism="Drosophila melanogaster"
	/mol_type="genomic DNA"
	/db_xref="taxon:7227"
	/clone="BACR32C19"
	/clone_lib="RPC1-98"

[illegible]

RESULT 3	
CNS0073W/c	
LOCUS	CNS0073W linear GSS 03-JUN-1999
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL066784
VERSION	AL066784.1 GI:4945247
KEYWORDS	GSS.
SOURCE	Drosophila melanogaster (fruit fly)
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 922)
REFERENCE	Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequences ;
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/The_BDGP_Drosophila_melanogaster_BAC_library prepared by Kazutoyo Osoegawa and Aaron Mammosee in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1. .922
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14D09"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN

Query Match 4.3%; Score 87.2; DB 9; Length 922;
Best Local Similarity 22.1%; Pred. No. 6.5e-13;
Matches 78; Conservative 145; Mismatches 130; Indels 0; Gaps 0;
Qy 930 CAACACAAACCAATCCCAAGATGCTGACTGAACCTAGACAAATAATACAACTGCACAA 989
Db 905 MMCM 846
Qy 990 ACCTGTCATGCCCCCTATACACTACCAATCTCTACTAACCAACCTCCAAACACAA 1049
Db 845 MYCMMAACMAACCMMA CMMAAAMAMMMMMMAAMMAACMAAMMAACMAAMMAAMMM 786
Qy 1050 CTTGAGCACTCTCTGACCACTTACAAAACACCAACCAATGACAAACACAGACCAAT 1109
Db 785 MMCM 726
Qy 1110 CACTGAAATGAGCAACGACGCTCCCTCCGATACCAACCTGCTCCAAACGGAATCC 1169
Db 725 MMMAAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAM 666
Qy 1170 CACCACGAAGACGACGACGAGCAAAAGAGGCGCCGACCAACGCGACCAACACGAC 1229
Db 665 AMMAAMMCCCAACAMCMCMCMCMCMMAAMMAAMMAAMMAAMMAAMMAAMMAAMMA 606
Qy 1230 AATGAGCACTTCCACGCTCTCCCCCACCCTCCGCTGCTGCTGCTGCTGCTGCTG 1282
Db 605 MMCM 553

RESULT 4

CNS006U0/C
LOCUS
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14N21 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION

AL065923.1 GI:4944891

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

COMMENT

Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 884)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1. .884
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14N21"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

Query Match 4.1%; Score 83.4; DB 9; Length 884;
Best Local Similarity 20.1%; Pred. No. 7.7e-12;
Matches 78; Conservative 171; Mismatches 138; Indels 1; Gaps 1;
Qy 855 CACCAACCAAGGGCTTTCATCAACAATGCCACCCACTCCCTCACCACAAACCAAGCAGCC 914
Db 881 MMVMVMVHTKKKKTTTHMMVMVMVMVMVMVMVMVMVMVMVMVMVMVMVMVMVMVM 822
Qy 915 ACAGCAAGGAGGAAACAAACCAATCCCAAGATGCTGTGACTGAACCTAGACAAAAA 974
Db 821 MMVM 762
Qy 975 TAACACAATGCAACACCGTCATGCGCCCTCATACACTACCAATCTCTACTACAA 1034
Db 761 MCMVM 702
Qy 1035 CACCTCAAAACACAACCTTACGACTCTCTGTGACCACTTACAAACACCAACCAATGACAA 1094
Db 701 MMAM-CHCMCMCHCM 643
Qy 1095 CACACAGACCAATCACTGAAATGAGCAAAACAGTGCCCTCGATATAACAACTGTC 1154
Db 642 CCCCCCMVMMAACAACAMCAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAMAAAM 583
Qy 1155 TCCAAACGGGAAATCCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1214
Db 582 CAMCTCACCACAAAHMAAHMAAHMAAHMAAHMAAHMAAHMAAHMAAHMAAHMA 523
Qy 1215 GGCAACAAACACGACAAATGAGCAATTC 1242
Db 522 KCBCMYBCCCCCCCCCAAMMAAMAAKCKC 495

RESULT 5

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

COMMENT

CNS00LO0
Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR32D23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
AL068607
AL068607.1 GI:4958689
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of

[illegible][illegible]


```

/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-17017"
/clone_lib="ISB1"
/note="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"

ORIGIN

Query Match      3.6%; Score 73.6; DB 9; Length 1965;
Best Local Similarity 48.8%; Pred. No. 6e-09;
Matches 199; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 875 CAACATGCCACCCACTCCCTCACCACCAACCCAGCCGACAGGAGGAACAACA 934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1185 CACCAACACCAACCAACCCACCCACCAACCCACCCACCCACCAACCCACACCC 1126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 935 CAACATGCCACCCACTCCCTCACCACCAACCCAGCCGACAGGAGGAACAACA 994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1125 ACACCAACCAACCCACCCACCCACCAACCCACCCACCCACCAACCAACCA 1066
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 995 CCATGCCCTTATACACTACCAATCTCTACTAACAACACCTTCCAAACACAACTTCA 1054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1065 CCCCCACACCCACACCCACCAACCAACCAACCCACCCACCCACCAACCAACCA 1006
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1055 GCATCTCTCTGCACATTACAAACACCAACCAATGACACACAGAGCAATCATCTG 1114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1005 CCCCACCAACACACCCACCCACCCACCAACCAACCAACCAACCAACCAACCA 946
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1115 AAATGAGCAACACAGTCCCTCGATACCAACCTGCTTCCACGAGGAAATCCACCA 1174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 945 CACACCAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 886
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1175 CAGCAAGAGCAGCAGCAGCAAAAGCCCGCCACCAACCGCAGCAACCAACGACAAATG 1234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 885 CAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 826
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1235 AGCATTTACAGTCTCTCCCGCCACCCGAGTGCAGTGCACAAATC 1282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 825 AAAAAAACAACCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCG 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
CNS006UU/c      939 bp      DNA      linear      GSS 03-JUN-1999
LOCUS           Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION      BAC14L09 of RPCI-98 library from Drosophila melanogaster (fruit
SOURCE          fly), genomic survey sequence.
ACCESSION       AL065953.1 GI:4944921
VERSION         GSS.
KEYWORDS        Drosophila melanogaster (fruit fly)
SOURCE          Drosophila melanogaster
ORGANISM        Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE       1 (bases 1 to 939)
AUTHORS         Genoscope.
TITLE           Direct Submission
JOURNAL         Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
COMMENT         Determination of this BAC-end sequence was carried out as part of a
                collaboration with the Berkeley Drosophila Genome Project (BDGP).
                The BDGP is constructing a physical map of the Drosophila
                melanogaster genome using these BACs. For further information
                please see http://www.fruitfly.org The BDGP Drosophila
                melanogaster BAC library was prepared by Kazutoyo Oosawa and
                Aaron Mammoss in Piter de Jong's laboratory in the Department of
                Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
                NY. The library is named RPCI-98 and was constructed by partial
                EcoRI digestion of drosophila DNA provided by the BDGP from the
                isogenic strain y2; cn bw sp, the same strain used for the BDGP's

```

```

Pl and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. 939
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BAC14L09"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

Query Match      3.5%; Score 72.6; DB 9; Length 939;
Best Local Similarity 33.8%; Pred. No. 9.2e-09;
Matches 124; Conservative 92; Mismatches 150; Indels 1; Gaps 1;

QY 795 CCTCGCAACATCCGGCTCCGGAGACGAGAGAGAGACGACGAGACGAGACGAG 854
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 865 MSSCMSACCCACRCMGMCAAGCGCCCRAGACRAGAGAGACGACGAGACGAG 806
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 855 CACCAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCACCAACGACGCC 914
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 805 VGSVVMACGAGSAGVSGVRCACACGCMVVCASCNMACAMMAMBMCMAMVASAASAA 746
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 915 ACAGCAAGAGGAAACAAACAAACCAATTCCTCCAGATGCTGTGACTGAACTAGACAAA 974
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 745 VAHVAGAVARGGAAAGCGRCGCAACRMCAACMCAACMAAMRMSACGSAACAMMMACAS 686
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 975 TAACAACTGCACACCGTCCATGCCCTCATACACTACCACTACCACTTCTACTAACAA 1034
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 685 VARMACAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 626
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1035 CACCTTCCAAACACCAACTTCA-GCACTCTCTCTGCACCACTTACAAAACACCAATGACA 1093
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 625 CCCCNCMSMAMACCACTCACTMYCKYCCCTCTTTTGGMCACAYCMYMAAMACA 566
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1094 ACACACAGAGCACAATCACTGAAATAGCAAAACAGTGCCCCCTCGATAACAACCCCTGC 1153
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 565 CMACMCCCTATCMMAAACACAAAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 506
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1154 CTCCAC 1160
    : : : : :
Db 505 MCCGCC 499

RESULT 12
CG754239/c      1787 bp      DNA      linear      GSS 24-OCT-2003
LOCUS           P049-3-D12.ya Ppa EcoRI BAC library Pristionchus pacificus genomic,
DEFINITION      genomic survey sequence.
ACCESSION       CG754239
VERSION         CG754239.1 GI:37979530
KEYWORDS        Pristionchus pacificus
SOURCE          Pristionchus pacificus
ORGANISM        Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
                Neodiplogasteridae; Pristionchus.
REFERENCE       1 (bases 1 to 1787)
AUTHORS         Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
                Buntjer,J., van der Meulen,M. and Sommer,R.J.
                An integrated physical and genetic map of the nematode Pristionchus
                pacificus
                Mol. Genet. Genomics 269 (5), 715-722 (2003)
TITLE           Evolutionary Biology
JOURNAL         Max-Planck-Institute for Developmental Biology
MEDLINE         Spemannstr. 37-39, Tuebingen D-72076, Germany
PUBMED          Tel: 00497071601371
                Fax: 00497071601498
                COMMENT

```


Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.

FEATURES

source
1. .1787
Location/Qualifiers
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of
the genomic DNA with EcoRI and cloning into the BAC
vector."

ORIGIN

Query Match 3.5%; Score 72.2; DB 9; Length 1787;
Best Local Similarity 48.9%; Pred. No. 1.5e-08;
Matches 194; Conservative 0; Mismatches 203; Indels 0; Gaps 0;
QY 883 CCACCACTCCCTCCACCAACAGCAGCGCCACAGGAGGAGGAAACACACAAACCAT 942
D 942
D 758 CCACCAACCATATCACACACACACACACACCTCCCATCAACACAAACACACACACCA 699
QY 943 TCCCAAGTGTCTGACTGACTGACCAAAATACACAACTGCACACCGTCCATGCC 1002
D 1002
D 698 ACACCAAGAGACGACCTAACCCACACACACCCGCCCAAAAAAAGCCACCCCCC 639
QY 1003 CCTATAACACTPACACACATCTCTACTAACACACTCTCAACACAACTTCAGCACTCTC 1062
D 1062
D 638 AACACCCCCACACACACACCAAAACAAACCAACACACACATACACACACACAC 579
QY 1063 TGTGACCATTTACAAACACACCAATGACACACACAGAGACACATCACTGAAATGAG 1122
D 1122
D 578 TCAACACATCAACAAATACACACACACCCCATACAACTTCCACCCACACCCCCAC 519
QY 1123 CAAACAGTGGCCCTCGATACACACCTGCTCCACAGGGAATCCACACACACAAAG 1182
D 1182
D 518 AACACACAAACAAACACACACACCCCAACCCCAACACACACACACATCCAC 459
QY 1183 AGCACACAGCAGCAAAAGGCGCCGCAACAGGCAACCAACACGACCAATGAGCATTT 1242
D 1242
D 458 ACCACATCCACACACACACACACACCCGCAACCAACCAACCAACCAACCAACCA 399
QY 1243 ACCAGTCTCCCCACACCCCGAGCTCGACTGCACAA 1279
D 1279
D 398 ACAACCCACCCAC 362

RESULT 13

AG131986
LOCUS Pan troglodytes DNA, 1626 bp DNA linear GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-144D13.R, genomic survey sequence.
ACCESSION AG131986
VERSION AG131986.1 GI:16661664
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
Submitted (02-AUG-2001) Asso Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Shirohoro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpansegsc.riken.go.jp, URL: http://hgp.sec.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end

COMMENT

was generated during the R&D process and may have higher chance of
clone tracking errors.

PRIMERS

Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
Location/Qualifiers
1. .1626
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-144D13.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES

source

ORIGIN

Query Match 3.5%; Score 71.8; DB 9; Length 1626;
Best Local Similarity 39.7%; Pred. No. 1.8e-08;
Matches 238; Conservative 0; Mismatches 362; Indels 0; Gaps 0;
QY 647 AAGATACAAATCTACAAAGAACCAACATGTCTCGTCCAAATACCTCCACCATGTC 706
D 706
D 156 AAANNAANANNCNCANAAAAACNANAAACAAANNNNNNNNCCNCCNCAANNACCCC 215
QY 707 CCACGCGCGTCCGAGATCAAACTCAACAGCACCCCAACTGATGCGCCACCAACTCAATA 766
D 766
D 216 CCACANNAACNNNAACCAANNAACCCNCCACCCNCCNNAACCAACCAAAAAA 275
QY 767 CCACGAGCCCAAGCATGATGAGGACCTCGCAACATCGGGTCCAGGTCGCGAGAAC 826
D 826
D 276 NANCNCCCAAAACAAAAAANNCANAAAAAANNAANNNNCAANAAAAANNCN 335
QY 827 GAGAACCCCAACAACTTCTGATGGGTCAACAAGCAAGGGGTTTCATCAACAATGCCAC 886
D 886
D 336 NNNNNCCCCNNNNNNNNAAAAAANAAAAAANAAAAAACCACCCCAAAAAA 395
QY 887 CCACCTCCCTCACCAACCAAGCAGCCACACAGGAGGAGAAACAAACCAACCACTTCCC 946
D 946
D 396 CCCAACCCCAACACCCNCCCNCAACAAAAAANCCNNCAANNAANNAACCAACCA 455
QY 947 AAGATGCTGTGACTGAACCTAGACAAAAAATAACACAACTGCACAAACCGTCCATGCCCTC 1006
D 1006
D 456 CACANACAAAAACAAAAANCCCAACCAACAAAAAACAACCCCCCAACACACAAAC 515
QY 1007 ATAACACTACCAATCTCTACTAACACACCTCCAAACACAACTTCAGCACTCTCTGTG 1066
D 1066
D 516 NACACACACCCACNANNAACCAAAAAACAAACCCCAACANNAACANWACACCCCA 575
QY 1067 CACCAATTACAAAAACACCAACCAATGACAAACACAGAGACCAATCACTGAAATGAGCAAA 1126
D 1126
D 576 CACAAACCCCAACACCCCCCAACAAAAAANAAAAAACCACCCCAACCAACCAACAA 635
QY 1127 CCAGTGGCCCTCGATACAAACCTCGCTCCCAACGGGAAATCCCAACAGCAAGAGCA 1186
D 1186
D 636 CCCCAACCCCCNCCCAACACACACNCCCAACCAACCAACCAACCAACCAACCA 695
QY 1187 CCAGCAGCAAAAAAGGCGCCCAACAGGCAACCAACCAACCAACCAACCAACCA 1246
D 1246
D 696 ACCCNCAACCAACCAACCCCAACCAACCAACCAACCAACCAACCAACCAACCA 755

RESULT 14

AG382036/c
LOCUS Mus musculus molossinus DNA, 1455 bp DNA linear GSS 03-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone: MSMg01-192F03.TJ, genomic survey
sequence.
ACCESSION AG382036
VERSION AG382036.1 GI:47993241
KEYWORDS GSS.
SOURCE Mus musculus molossinus


```
Db      872  AMAMAMAMACCMAMCCACCMCMCACCCCMCMCMAMCMCMCCCMCMAMCCCCCCCC 931
Qy      1042 AAACACAACTTCAGCACTCTCTGTGACCACTTACAAAAACACCACCAATGACAAACACACAG 1101
Db      932  ACCACMCCACCCCMCACCCCCCCCMCMCCCMCMCMAMMAACCMACMCMACCCCMMAC 991
Qy      1102 AGCACAATCACTGAAAAATGAGCAAAACCAAGTGCCCCCTCGATTAACAAACCTGGCTCCAACG 1161
Db      992  MCCACMCCCMCAAAAAAACCCMCHACMMCMCCACACACACAMCCCCCMCMACAAAMA 1051
Qy      1162 GGAAATCCCAACACAGCAAGAGCACCAGCAGCAAAAAAGGCCCGCCAC 1211
Db      1052 MAMAMACAAAMAMAMAMAAAAAACAMMAMMMAMCCAMMAMCCMMACAM 1101
```

Search completed: August 1, 2005, 10:09:09
Job time : 6682.18 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:35:29 ; Search time 327.71 Seconds
(without alignments)
10215.808 Million cell updates/sec

Title: US-10-066-506A-5
Perfect score: 2046
Sequence: 1 atgaagaccacatgttctct.....tcactaataacattggtga 2046

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1841.2	90.0	2164	3	US-08-760-615-3
2	1841.2	90.0	11460	4	US-09-336-910A-1
3	1510	73.8	2247	3	US-08-760-615-5
4	131.2	6.4	2172	3	US-08-760-615-1
5	131.2	6.4	2298	4	US-09-650-086A-1
C 6	74.6	3.6	152132	4	US-09-949-016-13845
C 7	74.6	3.6	152145	4	US-09-949-016-12371
8	60.4	3.0	34230	4	US-09-949-016-12052
9	60.4	3.0	128470	4	US-09-949-016-13765
10	57.6	2.8	5511	3	US-08-928-361B-2
11	57.6	2.8	5511	4	US-09-588-995A-2
12	57.6	2.8	7334	3	US-08-928-361B-1
13	57.6	2.8	7334	4	US-09-588-995A-1
C 14	54.2	2.6	30656	4	US-09-949-016-14613
C 15	54	2.6	522	4	US-09-949-016-103758
16	53.6	2.6	5163	3	US-08-700-651-1
17	53.6	2.6	5163	3	US-08-928-361B-4
18	53.6	2.6	5163	4	US-09-588-995A-4
19	53.6	2.6	5318	3	US-08-700-651-2
20	53.6	2.6	5318	3	US-08-928-361B-3
21	53.6	2.6	5318	4	US-09-588-995A-3
C 22	52.6	2.6	271134	4	US-09-949-016-12705
C 23	52.6	2.6	305491	4	US-09-949-016-17550
C 24	52.4	2.6	39154	4	US-09-949-016-12384
C 25	52.4	2.6	39154	4	US-09-949-016-12801
C 26	52.4	2.6	39443	4	US-09-949-016-14326
C 27	52.4	2.6	39443	4	US-09-949-016-14327

28	51.6	2.5	114793	4	US-10-148-806-3	Sequence 3, Appli
C 29	51	2.5	57280	4	US-09-949-016-11796	Sequence 11796, A
C 30	51	2.5	57280	4	US-09-949-016-12843	Sequence 12843, A
C 31	51	2.5	57280	4	US-09-949-016-12844	Sequence 12844, A
C 32	51	2.5	57280	4	US-09-949-016-12846	Sequence 12846, A
C 33	51	2.5	57280	4	US-09-949-016-13542	Sequence 13542, A
C 34	51	2.5	57280	4	US-09-949-016-13543	Sequence 13543, A
C 35	51	2.5	57280	4	US-09-949-016-13544	Sequence 13544, A
C 36	51	2.5	57280	4	US-09-949-016-13545	Sequence 13545, A
C 37	51	2.5	57280	4	US-09-949-016-14633	Sequence 14633, A
C 38	51	2.5	57280	4	US-09-949-016-14634	Sequence 14634, A
C 39	51	2.5	57280	4	US-09-949-016-14635	Sequence 14635, A
C 40	51	2.5	57280	4	US-09-949-016-14636	Sequence 14636, A
C 41	51	2.5	57280	4	US-09-949-016-14637	Sequence 14637, A
C 42	51	2.5	57280	4	US-09-949-016-14638	Sequence 14638, A
C 43	51	2.5	57280	4	US-09-949-016-14639	Sequence 14639, A
C 44	51	2.5	57280	4	US-09-949-016-14640	Sequence 14640, A
C 45	48.8	2.4	60990	4	US-09-949-016-14080	Sequence 14080, A

ALIGNMENTS

RESULT 1
US-08-760-615-3
; Sequence 3, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760.615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Heron, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2164 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Musoke
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 119..2164
US-08-760-615-3

Query Match									
Best Local Similarity 90.0%; Score 1841.2; DB 3; Length 2164;									
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;									
QY	1	ATGAAGACACATGTTTCCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC	60	QY	1081	ACCACCAATGACAAACACACAGAGCACAAATCACTGAAATGAGCAAAACCAAGTCCCCCTCG	1140		
DB	119	ATGAAGACACATGTTTCCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC	178	DB	1199	ACCACCAATGACAAACACACAGAGCACAAATCACTGAAATGAGCAAAACCAAGTCCCCCTCG	1258		
QY	61	ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT	120	QY	1141	ATAACAACCCCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCACACAGCAAAAAA	1200		
DB	179	ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT	238	DB	1259	ATAACAACCCCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCACACAGCAAAAAA	1318		
QY	121	CTCCAGAGACAGAGAGCTCCATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT	180	QY	1201	GGCCCGGCCAACAGCGCACCAAAACAGCAAAATGAGCAATTTCCACCACTCTCCCCCACC	1260		
DB	239	CTCCAGAGACAGAGAGAGCTCCATCTGATGGATTTCACACTGAGTGGGCAAAAAGTTGCT	298	DB	1319	GGCCCGGCCAACAGCGCACCAAAACAGCAAAATGAGCAATTTCCACCACTCTCCCCCACC	1378		
QY	181	GATTCCTCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAT	240	QY	1261	CCCAGCTCGACTGCACAAACATCTTGTATATTTTCAGAGAAAGCGATCGATTTCTCGAAA	1320		
DB	299	GATTCCTCTTTGGAGGCATCCAAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAT	358	DB	1379	CCCAGCTCGACTGCACAAACATCTTGTATATTTTCAGAGAAAGCGAGTATCTCTGGAGG	1438		
QY	241	GTGAGTACACAGAGGGGAGGAGCCAAAACATGCTACAATATAAGTGTAAACGGATCCC	300	QY	1321	GAAGTGATATATTTCCCGTTTTTAGATGGGTTAATAAATACTGAAATTTGATTTGATCCA	1380		
DB	359	GTGAGTACACAGAGGGGAGGAGCCAAAACATGCTACAATATAAGTGTAAACGGATCCC	418	DB	1439	GAAGTGACATGTTTCCCTTTTCTGGATGGGTTAATAATGCTCCAATTTGATTTGACCCA	1498		
QY	301	TCCTGAAAAATCCTTCTGTAGATCCTCTTACCAACATCCGTGACTATCCGAAATGCAAA	360	QY	1381	ATCCCAACACAGAAAAACAATCTTTTGATGAATCTCCCAGCTTTAATACTTCAACTAATAG	1440		
DB	419	TCCTGAAAAATCCTTCTGTAGATCCTCTTACCAACATCCGTGACTATCCGAAATGCAAA	478	DB	1499	GTTCCAATATCAAAAAACAATCTTTTGATGAATCTCTAGTTCTGGTCTCGGCTGAGGAA	1558		
QY	361	ACTATCCATATATTAAGGTCAAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG	420	QY	1441	GAACAACACACTCCCCGGAATATCAGTTTAACTTTCTCTTATTTTCTGATAAAAAATGGA	1500		
DB	479	ACTATCCATATATTAAGGTCAAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG	538	DB	1559	GATCAACATGCTCTCCCCCAATATATAGTTTAACTTTATCTTATTTTCTTAATAAATGAG	1618		
QY	421	GGAGCATTTTCTGTATGATCGCATGCTGCTCCCAACAATGTACCGAGGCAAACTTTC	480	QY	1501	GATCTGCTACTCTGGGGAAAAACGAGAATGATTTGTGATGACAGATTTGAGGATTTGGAGT	1560		
DB	539	GGAGCATTTTCTGTATGATCGCATGCTGCTCCCAACAATGTACCGAGGCAAACTTTC	598	DB	1619	AACACTGCTACTCTGGAGAAATGAGATGATTTGTGATGACAGATTTAGAAATTTGGAGC	1678		
QY	481	ACTGAAGGGAACATAGCAGCTATGATTTGTCAATAAGACAGTGCACAAAATGATTTTCTCG	540	QY	1561	GTGAGAGGACGATTTTGGCGGACGGCTTAGCTGGATACCAATTTTTTGGCCCTCGGAATC	1620		
DB	599	ACTGAAGGGAACATAGCAGCTATGATTTGTCAATAAGACAGTGCACAAAATGATTTTCTCG	658	DB	1679	GTTCAGGAGGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTTTGGCCCTGGAAAT	1738		
QY	541	CGGCAAGGACAAAGGTACCGTCAATAGATCTGACTTCTACTAATAAATATGGACAAGT	600	QY	1621	GAAGCATCTATCTGCGGTTTTAATCAAAAAATCAGAAACAAATTTAGTTTCTAGTTGAGG	1680		
DB	659	CGGCAAGGACAAAGGTACCGTCAATAGATCTGACTTCTACTAATAAATATGGACAAGT	718	DB	1739	GAAGCATTTTACACTGCTGTTTTAATTAATAATCAAAAAAATTTGGTCTGCAGGTTGAGG	1798		
QY	601	AGTAACGGAAACCAACGAATGACACTGGATGTTTTCGGGCTCTTCAAGATACAAATCT	660	QY	1681	CGCTTAGCTAATCAAACTGCTTAAATCCTTTGGAGCTCTTGTAAAGGTCACAAACCGAGAA	1740		
DB	719	AGTAACGGAAACCAACGAATGACACTGGATGTTTTCGGGCTCTTCAAGATACAAATCT	778	DB	1799	CGTCTAGCCAAATCAAACTGCCAAATCCTTTGGAATCTTATTGAGAGTCACAACTGAGGA	1858		
QY	661	ACAAAGAACCAACATGTGCTCCGTCGAAATACCTCCACCACTGCCACAGCCGCTCGG	720	QY	1741	AGGACATTTTCTTAAATCAATAGCGATCAATTTGACTTTTTTGTCTTACGAGGTGGGCGGA	1800		
DB	779	ACAAAGAACCAACATGTGCTCCGTCGAAATACCTCCACCACTGCCACAGCCGCTCGG	838	DB	1859	AGAACATTTCTTTAATCAATAGACATGCTATTGACTTTTCTACTCAAGATGGGAGGA	1918		
QY	721	GAGATCAAACTCAAGACACCCAACTGATGCAACCAACTCAATACAGGACCCCAAGC	780	QY	1801	ACATGCAAGTGTAGGACCTGATTGTCATAGGAATAGAGATCTATCTAAAAATATC	1860		
DB	839	GAGATCAAACTCAAGACACCCAACTGATGCAACCAACTCAATACAGGACCCCAAGC	898	DB	1919	ACATGCAAGTGTCTTGGACCTGATTGTCATCGGGATAGAGACTTGTCCAAAAATATT	1978		
QY	781	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACA	840	QY	1861	TCAGAACAAATTCGACAAAAATCAGAAAGGATGAAACAAAAGGAGGAACTTGGCTGGGCTCTA	1920		
DB	899	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCCAACA	958	DB	1979	TCAGAGCAAAATGACCAAAATTAAGAGGACGAAACAAAAGAGGGGACTGGTTGGGCTCTG	2038		
QY	841	ACTTCTGATGCGGTCAACAGCAAGGCTTTTCATCAACAATGCCACCACTCCCTCACCA	900	QY	1921	GGTGCAAAATGGTGACATCTGACTGGGGTGTCTTCAACAAATTTGGGCACTCTGTACTA	1980		
DB	959	ACTTCTGATGCGGTCAACAGCAAGGCTTTTCATCAACAATGCCACCACTCCCTCACCA	1018	DB	2039	GGTGGTAAATGGTGGAATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTGCTACTA	2098		
QY	901	CAACCAAGCACGCCACAGAGGAGGAAACAAACAAACCAATTTCCCAAGATGCTGTGACT	960	QY	1981	TTATCTATAGCTGTCTTGATTTGCTCTGTCTGTATCTGTCTGATCTTCTCACTAAATACATT	2040		
DB	1019	CAACCAAGCACGCCACAGAGGAGGAAACAAACAAACCAATTTCCCAAGATGCTGTGACT	1078	DB	2099	TTATCCATAGCTGCTTGAATGCTCTATCTCTGTATTGTCGTATTCTTTACTAAATATATC	2158		
QY	-			QY	2041	GGATGA 2046			
DB				DB	2159	GGATAA 2164			

RESULT 2

US-09-336-910A-1
; Sequence 1, Application US/09336910A
; Patent No. 6517842
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143
; CURRENT APPLICATION NUMBER: US/09/336,910A
; CURRENT FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: US 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
; ORGANISM: Marburg Virus
US-09-336-910A-1

Query Match 90.0%; Score 1841.2; DB 4; Length 11460;

Best Local Similarity 93.7%; Pred. No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

Qy	1	ATGAAGACCAATGTTCCCTTATCAGTCTTATCTTAAATTCAGGACACAAAAATCTCCC	60
Db	5940	ATGAAGACCAATGTTCCCTTATCAGTCTTATCTTAAATTCAGGACACAAAAATCTCCC	5999
Qy	61	ATTTAGAGATGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT	120
Db	6000	ATTTAGAGATGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT	6059
Qy	121	CTCAGAGACAGAGAGCTGATCTGATGGATTCACATGAGTGGGCAAAAGTTGCT	180
Db	6060	CTCAGAGACAGAGAGCTGATCTGATGGATTCACATGAGTGGGCAAAAGTTGCT	6119
Qy	181	GATTCCTCTTTGGAGCATCAAGGATGGGCTTTTCAGACAGGTGTACCTCCCAAGAAAT	240
Db	6120	GATTCCTCTTTGGAGCATCAAGGATGGGCTTTTCAGACAGGTGTACCTCCCAAGAAAT	6179
Qy	241	GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAACGGATCCC	300
Db	6180	GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAACGGATCCC	6239
Qy	301	TCTGAAAAATCTTGCTGTTAGATCCTCCTACCAACATCCGTGACTATCCGAAATGCAAA	360
Db	6240	TCTGAAAAATCTTGCTGTTAGATCCTCCTACCAACATCCGTGACTATCCGAAATGCAAA	6299
Qy	361	ACTATCCATCATATTCRAGGTCAAAACCTCATGACAGGGGATGCGCTTCATTTATGG	420
Db	6300	ACTATCCATCATATTCRAGGTCAAAACCTCATGACAGGGGATGCGCTTCATTTATGG	6359
Qy	421	GGAGCATTTTTCTGTATGATGCTATGCTCCCAACAAATGTACCGAGGCAAAAGTCTTC	480
Db	6360	GGAGCATTTTTCTGTATGATGCTATGCTCCCAACAAATGTACCGAGGCAAAAGTCTTC	6419
Qy	481	ACTGAAGGAAACATAGAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCG	540
Db	6420	ACTGAAGGAAACATAGAGCTATGATTTGTCATTAAGACAGTGCACAAAATGATTTTCG	6479
Qy	541	CGGCAAGACAGGCTACCGTCATATGATCTGACTTCTACTAATAATATTGGACAGT	600
Db	6480	CGGCAAGACAGGCTACCGTCATATGATCTGACTTCTACTAATAATATTGGACAGT	6539
Qy	601	AGTAACGGAACGCAAAACGAATGACATGGATGTTTTCGGGCTCTTCAAGAAATACAAATCT	660
Db	6540	AGTAACGGAACGCAAAACGAATGACATGGATGTTTTCGGGCTCTTCAAGAAATACAAATCT	6599

Qy	561	ACAAAGAACCAAAACATGCTGCTCCGTCCAAAATACCTCCACACTGCCACAGCCCGTCCG	720
Db	6600	ACAAAGAACCAAAACATGCTGCTCCGTCCAAAATACCTCCACACTGCCACAGCCCGTCCG	6659
Qy	721	GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC	780
Db	6660	GAGATCAAACTCAACAGCACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC	6719
Qy	781	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCACACA	840
Db	6720	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCACACA	6779
Qy	841	ACTTCTGATGGGTCAACAGCAAGGGCTTTTCATCAACATGCCACCACTCCCTCACA	900
Db	6780	ACTTCTGATGGGTCAACAGCAAGGGCTTTTCATCAACATGCCACCACTCCCTCACA	6839
Qy	901	CAACCAAGCAGCCACACAGAGGAGGAGAAACAAACAAACCAATTCCTCAAGATGCTGTACT	960
Db	6840	CAACCAAGCAGCCACACAGAGGAGGAGAAACAAACAAACCAATTCCTCAAGATGCTGTACT	6899
Qy	961	GACTAGACAAAAATAACACAACTGCAACCGTCCATGCCCTCTCATAACTACACACA	1020
Db	6900	GACTAGACAAAAATAACACAACTGCAACCGTCCATGCCCTCTCATAACTACACACA	6959
Qy	1021	ATCTTACTTAACAACACTCCAAAACAACTTTCAGCACTCTCTGACCACTTACAAAC	1080
Db	6960	ATCTTACTTAACAACACTCCAAAACAACTTTCAGCACTCTCTGACCACTTACAAAC	7019
Qy	1081	ACCACCAATGACAAACACACAGAGCACAATCACTCAAAATGAGCAAAACAGTCCCTCG	1140
Db	7020	ACCACCAATGACAAACACACAGAGCACAATCACTCAAAATGAGCAAAACAGTCCCTCG	7079
Qy	1141	ATAACAAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGACACAGCAAAAAA	1200
Db	7080	ATAACAAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGACACAGCAAAAAA	7139
Qy	1201	GGCCCCGCCACAAACGGGACCAACACAGCAAAATGAGCAATTCACAGTCTCCCTCCAC	1260
Db	7140	GGCCCCGCCACAAACGGGACCAACACAGCAAAATGAGCAATTCACAGTCTCCCTCCAC	7199
Qy	1261	CCAGCTCGACTGACACAACTCTTGTATATTCAGAGAAAGCGATCCATTTCTGGAAA	1320
Db	7200	CCAGCTCGACTGACACAACTCTTGTATATTCAGAGAAAGCGAGTATCTCTGGAGG	7259
Qy	1321	GAAGGTGATATATTTCCGTTTTTAGATGGGTAAATAACTGAAATTTGATTTGATCCA	1380
Db	7260	GAAGGTGATATATTTCCGTTTTTAGATGGGTAAATAACTGAAATTTGATTTGATCCA	7319
Qy	1381	ATCCAAACACAGAAACAACTTTTGTATGATGAAATCTCCAGCTTTAATACTTCAACTAATGAG	1440
Db	7320	ATCCAAACACAGAAACAACTTTTGTATGATGAAATCTCTAGTTCCTGAGTTCCTGAGGAA	7379
Qy	1441	GAACACACACTCCCGGAAATATCAGTTTAACTTTCTTATTTTCTGATAAATAATGGA	1500
Db	7380	GATCAACATGCTCCCAATATTTAGTTTAACTTTTCTTATTTTCTTAATATAATGAG	7439
Qy	1501	GATAGTCTACTCTGCGGAAACGAGAAATGATGATGATGATGATGATGATGATGATGATG	1560
Db	7440	AACACTGCTACTCTGCGGAAATGAGAAATGATGATGATGATGATGATGATGATGATG	7499
Qy	1561	GTGAGAGGAGCAGATTTGGCGGAGGGCTTAGCTGATACCAATTTTGGGCTGGAATC	1620
Db	7500	GTTCAGGAGGATGACTGCGGCGAGGGCTCAGTTGGATACCGTTTTTGGGCTGGAAT	7559
Qy	1621	GAGGACTCTATCTGCGGTTTAAATCAAAATCAGAACTTAACTTGTAGTGGTGGG	1680
Db	7560	GAGGACTCTATCTGCTGTTTTTAAATTAATAATCAAAATTTGGTCTGAGGTTGAGG	7619
Qy	1681	CGCTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGGTCAACACCGAGAA	1740
Db	7620	CGCTTAGCTAATCAAACTGCTCAAACTCTTGGAACTCTTATTTAGAGATCACAACCTGAGNA	7679

Db 1057 GAGCACAATGGNACCGACCCCAACACACACACGACGACGCTCCTCAACAATACTAATACA 1116
Qy 1021 ATCTCTACTAACAACAACCTCCAAACACAACTTTCAGCACTCTCTCTGACCACTTACAAAAC 1080
Db 1117 ACTCCCACTATAACACTCTCAAGTACAACTCTAGTACTCTCTCCCTCCCAACCGCAAC 1176
Qy 1081 ACCACCATGACACACACAGACACATCACTGAATAGACAAACGAGTGCCCTCG 1140
Db 1177 ATCAACCAATATGATACAAACGTCGCAACTAGCAGAAAGCAGCAACCAATGCTCAGTTG 1236
Qy 1141 ATAACAACCCCTCCCTCCAAACGGGAAATCCCAACAGCAAGAGAGACAGCAAGAAAA 1200
Db 1237 AACACAATCTAGATCCAAAGAGAAATCCCAACAGCAAGAGAGACAGCAAGCAAC 1296
Qy 1201 GGGCCCGCCAAACGGGACCAACACAGCAAAATAGCATTTTCAACAGTCTCCGCCAC 1260
Db 1297 AACATCATCATGACATCAGATATACAAAGCAACACCCCAAAATTTCTCTCCGAT 1356
Qy 1261 CCCAGCTGACGACACATCTGTATATTTTCAAGAGAGCGATCGATTTCTGGAA 1320
Db 1357 TCTAGTCGCAACCCGCCCTCTATATCTTTAGAAAGAACGAGCATTTCTCTGGAA 1416
Qy 1321 GAAGTGATATATTCCTGTTTGTAGTGGGTTAATAAATCTGAAATTTGATTTGATCCA 1380
Db 1417 GAAGTGATATATTCCTGTTTGTAGTGGGTTAATAAATCTGAAATTTGATTTGATCCA 1476
Qy 1381 ATCCAAACACAGAAACAACTTTTGATGAATCTCCAGCTTTAATACCTCAACTAATGAG 1440
Db 1477 ATCCAAACACAGAAACAACTTTTGATGAATCTCCAGCTTTAATACCTCAACTAATGAG 1536
Qy 1441 GAACACACACTCCCGCCGAATATCAGTTTAACTTTCTTCTTATTTCTGTGATAAATGGA 1500
Db 1537 GAACACACACTCCCGCCGAATATCAGTTTAACTTTCTTCTTATTTCTGTGATAAATGGA 1596
Qy 1501 GATCTGCTACTCTCGGGGAAAAACAGAAATGATTTGATGACAGATTTGAGGATTTGGAGT 1560
Db 1597 GATCTGCTACTCTCGGGGAAAAACAGAAATGATTTGATGACAGATTTGAGGATTTGGAGT 1656
Qy 1561 GTGAGGAGGACGATTTTGGCGGACGGGTTAGCTGATACCAATTTTGGCCCTGGAATC 1620
Db 1657 GTGAGGAGGACGATTTTGGCGGACGGGTTAGCTGATACCAATTTTGGCCCTGGAATC 1716
Qy 1621 GAAGGACTCTATCTGCGGGTTTAAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG 1680
Db 1717 GAAGGACTCTATCTGCGGGTTTAAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG 1776
Qy 1681 CGTTAGCTTAATCAAACTGCTAAATCTTTGGAGCTCTTTGTTAAAGGTCACAAACCGAGAA 1740
Db 1777 CGTTAGCTTAATCAAACTGCTAAATCTTTGGAGCTCTTTGTTAAAGGTCACAAACCGAGAA 1836
Qy 1741 AGGACATTTCTTAATCAATAGGATGCAATTTGCTTTTGGCTTACGAGGTGGGCGGA 1800
Db 1837 AGGACATTTCTTAATCAATAGGATGCAATTTGCTTTTGGCTTACGAGGTGGGCGGA 1896
Qy 1801 ACATCAAGGTGCTAGGACCTGATTTGTTGATAGAAATAGAGATCTATCTAAATAATC 1860
Db 1897 ACATCAAGGTGCTAGGACCTGATTTGTTGATAGAAATAGAGATCTATCTAAATAATC 1956
Qy 1861 TCAGAAACAAATCGAAATCAGAAAGGATGAAACAAAGAGGAGAACTGGCTGGGCTCTA 1920
Db 1957 TCAGAAACAAATCGAAATCAGAAAGGATGAAACAAAGAGGAGAACTGGCTGGGCTCTA 2016
Qy 1921 GGTGGCAATGGTGACATCTGACCTGGGGTGTCTTCAACCAATTTGGGCATCTGCTACTA 1980
Db 2017 GGTGGCAATGGTGACATCTGACCTGGGGTGTCTTCAACCAATTTGGGCATCTGCTACTA 2076
Qy 1981 TTATCTATAGCTGTTCTGATGCTCTGCTGATATCTGCTGATCTTCACTAAATACATT 2040
Db 2077 TTATCTATAGCTGTTCTGATGCTCTGCTGATATCTGCTGATCTTCACTAAATACATT 2136
Qy 2041 GGATGA 2046
Db 2137 GGATGA 2142

RESULT 4
US-08-760-615-1
; Sequence 1, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Ebola virus
; STRAIN: Zaire
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2172
; OTHER INFORMATION: /product= "Glycoprotein"
US-08-760-615-1

Query Match 6.4%; Score 131.2; DB 3; Length 2172;
Best Local Similarity 61.9%; Pred. No. 7.4e-30;
Matches 208; Conservative 0; Mismatches 128; Indels 0; Gaps 0;
Qy 1555 TGGAGTGTGAGGAGGAGGATTTGGCGGACGGCTTAGCTGATACCAATTTTGGCCCT 1614
Db 1693 TGGAGTGTGAGGAGGAGGATTTGGCGGACGGCTTAGCTGATACCAATTTTGGCCCT 1752
Qy 1615 GGAATCGAAGGACTCTATCTACTGCGGTTTAAATCAAAATCAGAAATTTAGTTTGTAGG 1674
Db 1753 GCAGCCGAGGAAATTACATAGAGGGGCTAATGCACATCAAGATGGTTTAACTGTGGG 1812
Qy 1675 TTGAGGCGCTTAGCTAATCAAAATCGCTTAAATCTTGGAGCTCTTGTAAAGGTCACAAAC 1734
Db 1813 TTGAGACAGCTGGCGCAACGAGACGACTCAAGCTCTTCAACTGTTCTTGAGAGCCCAACT 1872
Qy 1735 GAGGAAGGACATTTCTTCTTAATCAATAGGATCAATTTGACTTTTCTTACGAGGTGG 1794
Db 1873 GAGCTAGCACCTTTTCAATCCTCAACCGTAAGCAATTGATTTCTTCTGCTGAGCGATGG 1932


```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14613
; LENGTH: 30656
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(30656)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14613

Query Match      2.6%; Score 54.2; DB 4; Length 30656;
Best Local Similarity 14.5%; Pred. No. 6.3e-05;
Matches 164; Conservative 0; Mismatches 966; Indels 0; Gaps 0;

QY 142 CATCTGATGGGATTCACACTGAGTGGCGAAAGTTGCTGATTCCTTTGGAGGCATCC 201
DB 4633 CAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4574

QY 202 AAGCGATGGCTTTTCAGGACAGGTGTACTCTCCAAAGAAATGTTGAGTACACAGAGGGGAG 261
DB 4573 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4514

QY 262 GAAGCAGAAACATGCTACAAATAAGTGTACCGGATCCCTCTGGAAATTCCTTGCTGTGA 321
DB 4513 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4454

QY 322 GATCTCTTACCAACATCGTGACTATCGAAATGCAAACTATCCATCATATTCAAGGT 381
DB 4453 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4394

QY 382 CAAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGGGGAGCATTTTTCCTGTATGAT 441
DB 4393 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4334

QY 442 CGCATTTGCCCTCCACAACTGTACCGAGCAAGTCTTCACTGAAGGGAAACATAGCAGCT 501
DB 4333 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4274

QY 502 ATGATTGTCAATAAGACAGTGCACAAATGATTTTCTCGCGCAAGGACAAAGGTACCGT 561
DB 4273 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4214

QY 562 CATATGAATCTGACTTCTACTAATAAATATGGCAAGTAGTAGCAAGCAACGAAACGAAT 621
DB 4213 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4154

QY 622 GACATGTGATGTTTCGGGCGCTCTTCAAGAAATCAATTTCTTACAAAGAAACCAAACTGTGCT 681
DB 4153 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4094

QY 682 CGTCCAAAATACCTCCACCATGCCCCACAGCCCGTCCGGAGATCAAACTCACAAGCACC 741
DB 4093 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 4034

QY 742 CCAACTGTATGCCCAACAACTCAATACCAAGCAGCAAGCAGTGTATGAGGACCTCGCA 801
DB 4033 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3974

QY 802 ACATCGGGCTCAGGGTCCGGAGAAACGAGAAACCCCAACAACTTCTGTATGCGGTACCAAG 861
```

```
DB 3973 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3914
QY 862 CAAGGGCTTTTCATCAACAATGCCACCCACTCCTCTCACCAACAACGACGACGCCACGAA 921
DB 3913 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3854
QY 922 GGAGGAAACACAAACCACTTCCAAAGATGCTGTGACTGAAGTACAGAAAAATACACA 981
DB 3853 NNNNNNAACCACTTACCACCAACCAACCAATCAGCCATCAGCCATCACCACCAACCA 3794
QY 982 ACTGCACAACCGTCCATGCCCTCATAAACACTACCAATCTCTTACTTAAACAACCTCC 1041
DB 3793 ATCACCAGCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 3734
QY 1042 AAACACAACTTTCAGCACTCTCTCTGCACCACTTACAAAAACCAACCAACCAACCA 1101
DB 3733 AGCACTTATCATCATCTTACCACCATCATCATCATCATCATCATCATCATCATCATCA 3674
QY 1102 AGCACAATCACTGAAATGAGCAAAACGATGCCCTCGATGATAACACCTTGCCTCAAG 1161
DB 3673 ACCACCACTTACCACCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 3614
QY 1162 GGAATCCCAACAGCAAGAGCAGCAGCAGCAAGGCGCCGCGCCACCAACGACCA 1221
DB 3613 ACCATCACCAATCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3554
QY 1222 AACAGCAAAATGAGCATTTTCCAGCATCTCTCCCTCCCTCCCTCCCTCCCTCCG 1271
DB 3553 ATCACCACCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 3504
```

```
RESULT 15
US-09-949-016-103758/c
; Sequence 103758, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 103758
; LENGTH: 522
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-103758
```

```
Query Match      2.6%; Score 54; DB 4; Length 522;
Best Local Similarity 48.7%; Pred. No. 4.1e-06;
Matches 147; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 969 CAAAAATAACACAACTGCAACCGTTCATGCCCTCATTAACACTACCAATCTCTAC 1028
DB 487 CATCACCAACCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 428
QY 1029 TAAACAACCTTCCAAACAACTTTCAGCATCTCTCTGCACCAATTAACAAACCAACCA 1088
DB 427 CACCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 368
QY 1089 TGACAAACACAGAGCAAACTCACTGAAATGAGCAACCAACCAACCAACCAACCA 1148
DB 367 CAGCAGCAACCAATCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 308
```

Qy	1149	CCTGCTC	CAACGGGAATCC	CACAGCAAGAGCACCAGCAGCAAAAGGCGCCGC	1208
Db	307	CATTAC	SACCATCTAC	CACCATCACCAGCACCAGCACCACCAACCAT	248
Qy	1209	CACAA	CGGCACCAAA	CACGCAAAATGAGCATTTTCAACCAAGTCCTCCCCCAGCCCGCAGCTC	1268
Db	247	CACAA	CGGCACCAAT	CACCACTACCATCATCACCAATATCACCGAC	188
Qy	1269	GA	1270		
Db	187	CA	186		

Search completed: August 1, 2005, 10:31:12
Job time : 331.71 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 16:57:35 ; Search time 1285.84 Seconds
(without alignments)
10298.647 Million cell updates/sec

Title: US-10-066-506A-5
Perfect score: 2046
Sequence: 1 atgaagaccacatgttctct.....tcactaaatcatgtgatga 2046

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	2046	100.0	2046	15	US-10-066-506A-5
2	1841.2	90.0	7778	20	US-10-491-121-30
3	1841.2	90.0	11460	16	US-10-267-322-1
4	1813.6	88.6	2051	15	US-10-066-506A-9
5	1767.6	86.4	7005	20	US-10-491-121-31
6	1767.6	86.4	8256	20	US-10-491-121-32
7	1692.4	82.7	2046	21	US-10-811-353-2

8	1692.4	82.7	19112	17	US-10-353-856-27	Sequence 27, Appl
9	1518	74.2	2046	15	US-10-066-506A-13	Sequence 13, Appl
10	1319.6	64.5	2046	15	US-10-066-506A-7	Sequence 7, Appl
11	1308	63.9	1841	15	US-10-066-506A-3	Sequence 3, Appl
12	995.6	48.7	6902	20	US-10-491-121-42	Sequence 42, Appl
13	546.6	26.7	2252	15	US-10-066-506A-1	Sequence 1, Appl
14	145	7.1	6940	20	US-10-491-121-15	Sequence 15, Appl
15	145	7.1	7073	22	US-10-860-878-2	Sequence 2, Appl
16	145	7.1	7082	20	US-10-491-121-13	Sequence 13, Appl
17	145	7.1	7087	20	US-10-491-121-14	Sequence 14, Appl
18	145	7.1	8221	20	US-10-491-121-27	Sequence 27, Appl
19	145	7.1	8338	20	US-10-491-121-26	Sequence 26, Appl
20	145	7.1	10783	20	US-10-491-121-25	Sequence 25, Appl
21	138.8	6.8	6885	20	US-10-491-121-18	Sequence 18, Appl
22	138.8	6.8	6889	20	US-10-491-121-19	Sequence 19, Appl
23	138.8	6.8	7002	20	US-10-491-121-16	Sequence 16, Appl
24	138.8	6.8	7003	22	US-10-860-878-1	Sequence 1, Appl
25	138.8	6.8	7023	20	US-10-491-121-21	Sequence 21, Appl
26	138.8	6.8	7036	20	US-10-491-121-17	Sequence 17, Appl
27	138.8	6.8	8146	20	US-10-491-121-20	Sequence 20, Appl
28	131.2	6.4	2039	15	US-10-066-506A-11	Sequence 11, Appl
29	131.2	6.4	2298	9	US-09-337-946A-1	Sequence 1, Appl
30	131.2	6.4	2298	17	US-10-384-976-1	Sequence 1, Appl
31	131.2	6.4	2298	18	US-10-326-795-1	Sequence 1, Appl
32	131.2	6.4	2298	22	US-10-696-633-1	Sequence 1, Appl
33	131.2	6.4	6467	20	US-10-491-121-10	Sequence 10, Appl
34	131.2	6.4	6561	20	US-10-491-121-4	Sequence 4, Appl
35	131.2	6.4	6624	20	US-10-491-121-3	Sequence 3, Appl
36	131.2	6.4	6914	20	US-10-491-121-9	Sequence 9, Appl
37	131.2	6.4	7154	20	US-10-491-121-1	Sequence 1, Appl
38	131.2	6.4	7188	20	US-10-491-121-2	Sequence 2, Appl
39	131.2	6.4	7272	22	US-10-860-878-4	Sequence 4, Appl
40	131.2	6.4	7285	22	US-10-860-878-3	Sequence 3, Appl
41	131.2	6.4	8199	20	US-10-491-121-29	Sequence 29, Appl
42	131.2	6.4	8439	20	US-10-491-121-28	Sequence 28, Appl
43	131.2	6.4	18959	17	US-10-353-856-19	Sequence 19, Appl
44	131.2	6.4	18959	17	US-10-353-856-37	Sequence 37, Appl
45	131.2	6.4	18959	17	US-10-353-856-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-10-066-506A-5
; Sequence 5, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh_Microsoft Word 6.0
; SEQ ID NO 5
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein 2
; OTHER INFORMATION: Marburg virus strain Raven Glycoprotein 2
US-10-066-506A-5

Query Match 100.0%; Score 2046; DB 15; Length 2046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGACCACATGTTTCCTTATCATGCTTATCTTATTAATTCAGGGACAAAATCTCCCC 60

|||||
1 ATGAAGACCACATGTTTCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 120
61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 120
121 CTCAGAGACAGAAAGACGTCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
121 CTCAGAGACAGAAAGACGTCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
181 GATTCCTTTGGAGGACATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
181 GATTCCTTTGGAGGACATCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA 240
241 GTTGAATACAGAGGGGGAGGAGCCAAAATCATCTCAATATAGTGAACGATCCC 300
241 GTTGAATACAGAGGGGGAGGAGCCAAAATCATCTCAATATAGTGAACGATCCC 300
301 TCTGGAATATCTTGTCTGTAGATCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
301 TCTGGAATATCTTGTCTGTAGATCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
361 ACTATCCATCATATTCGAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
361 ACTATCCATCATATTCGAAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
421 GGAGCATTTTCTGTATGATCGCATGCTCCCAACAAATGATCCGAGGGCAAGTCTTC 480
421 GGAGCATTTTCTGTATGATCGCATGCTCCCAACAAATGATCCGAGGGCAAGTCTTC 480
481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATAGACAGTGCACAAAATGATTTCTCG 540
481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATAGACAGTGCACAAAATGATTTCTCG 540
541 CGGCAAGGCAAGGGTACCGTCAATATGATCTGACTTCTACTAATAATATTTGGCAAGT 600
541 CGGCAAGGCAAGGGTACCGTCAATATGATCTGACTTCTACTAATAATATTTGGCAAGT 600
601 AGTAAGGGAACCAACGATGACACTGGATGTTTCGGGCTCTTCAGAAATACAAATCT 660
601 AGTAAGGGAACCAACGATGACACTGGATGTTTCGGGCTCTTCAGAAATACAAATCT 660
661 ACAAGAACCAACATGTCTCGTCCAAATACCTCCACACTGCCACAGCCCGTCCG 720
661 ACAAGAACCAACATGTCTCGTCCAAATACCTCCACACTGCCACAGCCCGTCCG 720
721 GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780
721 GAGATCAAACTCAAGACACCCCACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780
781 AGTGATGAGGACCTCGCAATCCGGCTCAGGGTCGGAGACGAGAACCCACACA 840
781 AGTGATGAGGACCTCGCAATCCGGCTCAGGGTCGGAGACGAGAACCCACACA 840
841 ACTTCTGATCGGTCAACGAAGGCTTTTCAACAAATGCCACCACTCCCTCACCA 900
841 ACTTCTGATCGGTCAACGAAGGCTTTTCAACAAATGCCACCACTCCCTCACCA 900
901 CAACGAAGCAGCCACAGCAAGGAGAAACAACAACCAATTCCTGAGTCTGTGACT 960
901 CAACGAAGCAGCCACAGCAAGGAGAAACAACAACCAATTCCTGAGTCTGTGACT 960
961 GAAGTACAAAATTAACAACACTGCAACCGTCCATGCCCCCTCATAACACTACCACA 1020
961 GAAGTACAAAATTAACAACACTGCAACCGTCCATGCCCCCTCATAACACTACCACA 1020
1021 ATCTCTACTAACAACACTCCAAAACAACCTTTCAGCACTCTCTCTGCACATTAACAAC 1080
1021 ATCTCTACTAACAACACTCCAAAACAACCTTTCAGCACTCTCTCTGCACATTAACAAC 1080
1081 ACCACCAATGACAACACAGAGCAAAATCACTGAAAAATGAGCAAAACCAAGTCCCCCTCG 1140
|||||

Db 1081 ACCACCAATGACAACACAGAGCACAAATCACTGAAAATAGCAAAACAGTGCCCCCTCG 1140
QY 1141 ATAACAACCCCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAACAGCAGCAAAAAA 1200
Db 1141 ATACACCCCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAACAGCAGCAAAAAA 1200
QY 1201 GGCCCGGCAACAACGGGACCAAAACAGCAAAATGAGCAATTTCAACAGTCTCTCCCCAC 1260
Db 1201 GGCCCGGCAACAACGGGACCAAAACAGCAAAATGAGCAATTTCAACAGTCTCTCCCCAC 1260
QY 1261 CCCAGCTCGATGACACAACTCTTGTATATTTCAAGAAAGAGCGATCGATTTCTGGAA 1320
Db 1261 CCCAGCTCGATGACACAACTCTTGTATATTTCAAGAAAGAGCGATCGATTTCTGGAA 1320
QY 1321 GAAGTGATATATTTCCCGTTTTAGATGGGTTAATAAATCTGAAATTTGATTTTCTGATCA 1380
Db 1321 GAAGTGATATATTTCCCGTTTTAGATGGGTTAATAAATCTGAAATTTGATTTTCTGATCA 1380
QY 1381 ATCCCAACACAGAAACAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1440
Db 1381 ATCCCAACACAGAAACAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1440
QY 1441 GAACACACACTCCCGGAAATATCAGTTTAATCTTCTTATTTTCTGATAAATAATGGA 1500
Db 1441 GAACACACACTCCCGGAAATATCAGTTTAATCTTCTTATTTTCTGATAAATAATGGA 1500
QY 1501 GATCTGCTACTCTCGGGGAAAAACAGAAATGATGTGATGAGAGTTGAGGATTTGGAGT 1560
Db 1501 GATCTGCTACTCTCGGGGAAAAACAGAAATGATGTGATGAGAGTTGAGGATTTGGAGT 1560
QY 1561 GTGAGGAGGACGATTTGGCGGACGGCTTAGCTGGATACCAATTTTGGCCCTCGAATC 1620
Db 1561 GTGAGGAGGACGATTTGGCGGACGGCTTAGCTGGATACCAATTTTGGCCCTCGAATC 1620
QY 1621 GAAGACTCTATCTACTGCGGTTTTAATCAAAATCAGAACTTTAGTTTGTAGGTTGAG 1680
Db 1621 GAAGACTCTATCTACTGCGGTTTTAATCAAAATCAGAACTTTAGTTTGTAGGTTGAG 1680
QY 1681 CGTTAGCTAATCAAACTGCTAAATCTTGGAGCTCTTGTAAAGGTCAACACCGAGAA 1740
Db 1681 CGTTAGCTAATCAAACTGCTAATCTTGGAGCTCTTGTAAAGGTCAACACCGAGAA 1740
QY 1741 AGGACATTTCTTAAATCAATAGGACATCAATGATCTTTTGTCTTACGAGTGGGCGGA 1800
Db 1741 AGGACATTTCTTAAATCAATAGGACATCAATGATCTTTTGTCTTACGAGTGGGCGGA 1800
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAATAATATC 1860
Db 1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAATAATATC 1860
QY 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAAGGAGGAACTGGCTGGGCTCTA 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAAACAAAAGGAGGAACTGGCTGGGCTCTA 1920
QY 1921 GGTGCAAAATGGTGACATCTGATGGGTTTCTTACCAATTTGGGCACTCTGCTACTA 1980
Db 1921 GGTGCAAAATGGTGACATCTGATGGGTTTCTTACCAATTTGGGCACTCTGCTACTA 1980
QY 1981 TTAATCTAGTCTGATGCTCTGCTGCTGATCTGCTGATCTTCACTAATAATACAT 2040
Db 1981 TTAATCTAGTCTGATGCTCTGCTGCTGATCTGCTGATCTTCACTAATAATACAT 2040
QY 2041 GGATGA 2046
Db 2041 GGATGA 2046

RESULT 2

US-10-491-121-30
; Sequence 30, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY

APPLICANT: YANG, ZHI-YONG
APPLICANT: SULLIVAN, NANCY
APPLICANT: SANCHEZ, ANTHONY
TITLE OF INVENTION: Development of a Preventive Vaccine for
FILE OF INVENTION: Filovirus Infection in Primates
FILE REFERENCE: NIH221.001NP
CURRENT APPLICATION NUMBER: US/10/491,121
CURRENT FILING DATE: 2004-03-26
PRIORITY APPLICATION NUMBER: PCT/US02/30251
PRIOR FILING DATE: 2002-09-24
PRIOR APPLICATION NUMBER: US 60/326476
PRIOR FILING DATE: 2001-10-01
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 7778
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Construct pVR1012 Marburg
US-10-491-121-30

Query Match 90.0%; Score 1841.2; DB 20; Length 7778;

Best Local Similarity 93.7%; Pred. No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY	1	ATGAAGACACATGTTCTTATCAGTCTTATCTTAAATTCAGGGACAAAAATCTCCC	60
DB	2034	ATGAAGACACATGTTCTTATCAGTCTTATCTTAAATTCAGGGACAAAAATCTCCC	2093
QY	61	ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGGAATCGGTATGTCGCGAACT	120
DB	2094	ATTTTAGAGTAGCTAGTAATAATCAACCCCAAAATGGAATCGGTATGTCGCGAACT	2153
QY	121	CTCCAGAGACAGAAAGAGTCTCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	180
DB	2154	CTCCAGAGACAGAAAGAGTCTCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT	2213
QY	181	GATTCCTCTTTGAGGCAATCCAAAGCGATGGGCTTTTCAGGACAGAGTGATCTCCCAAGAT	240
DB	2214	GATTCCTCTTTGAGGCAATCCAAAGCGATGGGCTTTTCAGGACAGAGTGATCTCCCAAGAT	2273
QY	241	GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATAGGTAAACGGATCCC	300
DB	2274	GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATAGGTAAACGGATCCC	2333
QY	301	TCTGGAATCTCTGCTGTAGATCCTCTACCAACATCCCGTACTATCCGAAATGCAAA	360
DB	2334	TCTGGAATCTCTGCTGTAGATCCTCTACCAACATCCCGTACTATCCGAAATGCAAA	2393
QY	361	ACTATCCATCATATTCAGAGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	420
DB	2394	ACTATCCATCATATTCAGAGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG	2453
QY	421	GGAGCATTTTCTGTATGATCGATGCTGCTCCACAAATGATGACAGGGGCAAGTCTTC	480
DB	2454	GGAGCATTTTCTGTATGATCGATGCTGCTCCACAAATGATGACAGGGGCAAGTCTTC	2513
QY	481	ACTGAAGGGAACATAGCAGTATGATGTCATAATAGACAGTGCACAAAATGATTTCTCG	540
DB	2514	ACTGAAGGGAACATAGCAGTATGATGTCATAATAGACAGTGCACAAAATGATTTCTCG	2573
QY	541	CGGCAAGGACAAAGGTACCGTATATGATCTGACTTCTTACTAATAAATATTTGGACAAGT	600
DB	2574	CGGCAAGGACAAAGGTACCGTATATGATCTGACTTCTTACTAATAAATATTTGGACAAGT	2633
QY	601	AGTAACGGAAACGAAACGAAATGACACTGATGTTTGGGGCTCTTCAAGATATCAATTTCT	660
DB	2634	AGTAACGGAAACGAAACGAAATGACACTGATGTTTGGGGCTCTTCAAGATATCAATTTCT	2693
QY	661	ACAAAGAACCAACATGTCCTCGTCCCAAAATACCTCCACCACTGCCACAGCCCGTCCG	720
DB	2694	ACAAAGAACCAACATGTCCTCGTCCCAAAATACCTCCACCACTGCCACAGCCCGTCCG	2753

QY	721	GAGATCAAACTCACAAGCAGCCCAACTGATGTCGACCAAACTCAATAACACGAGCCCAAGC	780
DB	2754	GAGATCAAACTCACAAGCAGCCCAACTGATGTCGACCAAACTCAATAACACGAGCCCAAGC	2813
QY	781	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA	840
DB	2814	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA	2873
QY	841	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCGCAACCCCTCCACCA	900
DB	2874	ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCGCAACCCCTCCACCA	2933
QY	901	CAACCAAGCAGCCACAGCAGAGGAGGAAACAACAACCACTTCCCAAGATGCTGTGACT	960
DB	2934	CAACCAAGCAGCCACAGCAGAGGAGGAAACAACAACCACTTCCCAAGATGCTGTGACT	2993
QY	961	GAACTAGACAAAAATAACACAACCTGCACAACCGTCCATGCCCTCTATAAACAATACCACA	1020
DB	2994	GAACTAGACAAAAATAACACAACCTGCACAACCGTCCATGCCCTCTATAAACAATACCACA	3053
QY	1021	ATCTCTACTAAACAACCTCCANAACAACCTTTCAGCACTCTCTCTGCAACCAATTAACAAC	1080
DB	3054	ATCTCTACTAAACAACCTCCANAACAACCTTTCAGCACTCTCTCTGCAACCAATTAACAAC	3113
QY	1081	ACCACCAATGACACACACAGACACCAATCACTGAAATGAGCAAAACAGTGCCTCCCTCG	1140
DB	3114	ACCACCAATGACACACACAGACACCAATCACTGAAATGAGCAAAACAGTGCCTCCCTCG	3173
QY	1141	ATAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGCAGCAAAAAA	1200
DB	3174	ATAACAACCTCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCAGCAGCAGCAAAAAA	3233
QY	1201	GGCCCGCCACACAGCGCACCAAAACGACAAATGAGCAATTTTCCAGCTCTCTCCCTCCAC	1260
DB	3234	GGCCCGCCACACAGCGCACCAAAACGACAAATGAGCAATTTTCCAGCTCTCTCCCTCCAC	3293
QY	1261	CCAGCTGCACTGACACCAATCTTGTATATTTTCAGAAAGAGCGATCGATTTTCTGGAA	1320
DB	3294	CCAGCTGCACTGACACCAATCTTGTATATTTTCAGAAAGAGCGATCGATTTTCTGGAG	3353
QY	1321	GAAAGTGATATATTTCCCGTTTTAGATGGGTAAATAAATACTGAAATTTGATTTTCATCA	1380
DB	3354	GAAAGTGATATATTTCCCGTTTTAGATGGGTAAATAAATACTGAAATTTGATTTTCATCA	3413
QY	1381	ATCCCAAAACACAGAAACAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG	1440
DB	3414	GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTTCCTGGTCCCTCGCTGAGGAA	3473
QY	1441	GAAACAACAATCTCCCGCGAATATCAAGTTTAACTTTCTCTTATTTTCTGATAAAAATGGA	1500
DB	3474	GATCAACAATGCTCCCGCAATATAGTTTAACTTTTATCTTATTTTCTTAATAATAATGAG	3533
QY	1501	GATCTGCTACTCTGCGGAAACGAGAAATGATTTGATGATGAGCTTTCAGGATTTGGAGT	1560
DB	3534	AACACTGCTACTCTGCGGAAATGAGAAATGATTTGATGATGAGCTTTCAGGATTTGGAGC	3593
QY	1561	GTGACGAGGAGCAGATTTGCGCGCAGGGCTTACGTGATACCAATTTTGGCCCTCGAATC	1620
DB	3594	GTTCAAGGAGGATGACCTGCGCGCAGGGCTCAGTTGGATACCGTTTTTGGCCCTCGAAT	3653
QY	1621	GAAGGACTCTATATCTGCGGTTTTTAATCAAAAAATCAAGAAATTAAGTTGTTGAGTTGAGG	1680
DB	3654	GAAGGACTCTATATCTGCGGTTTTTAATTAATAAATAAATAAATTTGCTGCAAGTTGAGG	3713
QY	1681	CGCTTAGCTAATCAAACTGCTAATCTTGGAGCTCTTGTAAAGGTCACACCCGAGGA	1740
DB	3714	CGCTTAGCTAATCAAACTGCTAATCTTGGAGCTCTTGTAAAGGTCACACCCGAGGA	3773
QY	1741	AGGACATTTCTCTTAATCAATAGGATGCAATTTGACTTTTGTCTTACGAGGTGGCGGGA	1800
DB	3774	AGGACATTTCTCTTAATCAATAGGATGCAATTTGACTTTTGTCTTACGAGGTGGCGGGA	3833

QY 1801 ACATCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTTAAATATC 1860
Db 3834 ACATCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTTCCAAATAT 3893
QY 1861 TCAGAACAAATCGACAAATCAGAAAGATGAAAGGAGGAACTGGCTGGGCTCTA 1920
Db 3894 TCAGAACAAATCGACAAATCAGAAAGATGAAAGGAGGAACTGGCTGGGCTCTG 3953
QY 1921 GGTGCAAAATGGTGACATCTGACTGGGCTGTTCTCACCAATTTGGGCATCTCTGCTACTA 1980
Db 3954 GGTGTAATGGTGACATCTGACTGGGCTGTTCTTACTACTGGGATTTTGGCTACTA 4013
QY 1981 TTATCTATAGCTGTTCTGATGCTCTGCTCTGCTATCTGCTGATCTTCACTAAATACAT 2040
Db 4014 TTATCCATAGCTGTTCTGATGCTCTGATCTCTGATTTTGTGCTATCTTTACTAAATATATC 4073
QY 2041 GGATGA 2046
Db 4074 GGATAA 4079

RESULT 3
US-10-267-322-1
; Sequence 1, Application US/10267322
; Publication No. US20030152590A1
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143D
; CURRENT APPLICATION NUMBER: US/10/267,322
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 09/336,910, and US 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
; ORGANISM: Marburg Virus
US-10-267-322-1

Query Match 90.0%; Score 1841.2; DB 16; Length 11460;
Best Local Similarity 93.7%; Pred. No. 0;
Matches 1918; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY 1 ATGAGACACATGTTTCCCTTATCAGTCTTATCTTAAATTCAGGGACAAATAATCTCCCC 60
Db 5940 ATGAGACACATGTTTCCCTTATCAGTCTTATCTTAAATTCAGGGACAAATAATCTCCCC 5999
QY 61 ATTTTAGATAGCTAGTAAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 6000 ATTTTAGATAGCTAGTAAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 6059
QY 121 CTCGAGACAGAGAACGCTCCATCTGATGGGATTCACATCGATGGGCAAAAATGTTGCT 180
Db 6060 CTCGAGACAGAGAACGCTCCATCTGATGGGATTCACATCGATGGGCAAAAATGTTGCT 6119
QY 181 GATTCCTCTTTGGAGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db 6120 GATTCCTCTTTGGAGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 6179
QY 241 GTTGAGTACAGAGGGGGAGGAGCAAAACATGCTACAAATATAGTGAACGATCC 300
Db 6180 GTTGAGTACAGAGGGGGAGGAGCAAAACATGCTACAAATATAGTGAACGATCC 6239
QY 301 TCTGGAATAATCTTGTGTTAGATCTCTTCAACATCCGTGATCTCCGAAATGCAAA 360
Db 7320 GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTTCTGGTGCCTCGGCTGAGGAA 7379

Db 6240 TCTGGAATAATCTTGTGTTAGATCTCTTCAACAAATCCGTGACTATCTTAAATGCAAA 6299
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db 6300 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATGG 6359
QY 421 GGAGCATTTTTTCTGTATGATCGCATTTGCCCTCCACAAATGTACCGAGGCAAAAGTCTTC 480
Db 6360 GGAGCATTTTTTCTGTATGATCGCATTTGCCCTCCACAAATGTACCGAGGCAAAAGTCTTC 6419
QY 481 ACTGAAGGAAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAATGATTTTCTCG 540
Db 6420 ACTGAAGGAAACATAGCAGCTATGATTTGTCAATAGACAGTGCACAAATGATTTTCTCG 6479
QY 541 CGGCAAGGAAACAGGGTACCGTCAATATGAATCTGACTTCTTACTAATAAATAATTTGCAAGT 600
Db 6480 CGGCAAGGAAACAGGGTACCGTCAATATGAATCTGACTTCTTACTAATAAATAATTTGCAAGT 6539
QY 601 AGTAACGGAAACGCAAAACGAATGACACTGGATGTTTGGCGCTCTTTCAAGAATACAAATCT 660
Db 6540 AGTAACGGAAACGCAAAACGAATGACACTGGATGTTTGGCGCTCTTTCAAGAATACAAATCT 6599
QY 661 ACAAGAACCAAAACATGTGCTCCGTCCAAATACCTCCACCACTGCCACAGCCCGCTCG 720
Db 6600 ACAAGAACCAAAACATGTGCTCCGTCCAAATACCTCCACCACTGCCACAGCCCGCTCG 6659
QY 721 GAGATCAAACTCAACAGCACCCCAACTGATGCACCAAACTCAATACCAAGGACCAAGC 780
Db 6660 GAGATCAAACTCAACAGCACCCCAACTGATGCACCAAACTCAATACCAAGGACCAAGC 6719
QY 781 AGTATGATGAGGACCTCGCAACATCTCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
Db 6720 AGTATGATGAGGACCTCGCAACATCTCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 6779
QY 841 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA 900
Db 6780 ACTTCTGATGCGGTCAACAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA 6839
QY 901 CAACCAAGCAGCCACAGCAGAGGAGGAAACAAACAAACCACTTCCCAAGATCTGTGACT 960
Db 6840 CAACCAAGCAGCCACAGCAGAGGAGGAGGAAACAAACAAACCACTTCCCAAGATCTGTGACT 6899
QY 961 GAACTAGACAAATAATACAACTGCAACACCGTCCATGCCCCCTCATAACTACCAACA 1020
Db 6900 GAACTAGACAAATAATACAACTGCAACACCGTCCATGCCCCCTCATAACTACCAACA 6959
QY 1021 ATCTCTACTACACACCTCCAAACACAACTTCAGCACTCTCTGCACTTACCACTTACAAAC 1080
Db 6960 ATCTCTACTACACACCTCCAAACACAACTTCAGCACTCTCTGCACTTACCACTTACAAAC 7019
QY 1081 ACCACCAATGACACACAGAGCAGCAATCACTGAAAAATGAGCAAAACAGTGCCCCCTCG 1140
Db 7020 ACCACCAATGACACACAGAGCAGCAATCACTGAAAAATGAGCAAAACAGTGCCCCCTCG 7079
QY 1141 ATAAACACCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAACAGCAGCAAAAAA 1200
Db 7080 ATAAACACCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGCAACAGCAGCAAAAAA 7139
QY 1201 GGGCCGGCCCAACAGGCAACCAACAGCAAAATGAGCAATTCACCACTGCTCCCCCACC 1260
Db 7140 GGGCCGGCCCAACAGGCAACCAACAGCAAAATGAGCAATTCACCACTGCTCCCCCACC 7199
QY 1261 CCCAGCTCGACTGCAACACATCTTGTATATTTTCAAGAAGAGCGATCGATTTCTTGGAAA 1320
Db 7200 CCCAGCTCGACTGCAACACATCTTGTATATTTTCAAGAAGAGCGATCGATTTCTTGGAGG 7259
QY 1321 GAAGGTGATATATTCCTGTTTTTAGATGGGTAAATAAATACTGAAATGATTTGATCA 1380
Db 7260 GAAGGTGATATATTCCTGTTTTTAGATGGGTAAATAAATACTGAAATGATTTGATCA 7319
QY 1381 ATCCCAACACAGAAACAACTTTTGAATGATGATGATGATGATGATGATGATGATGATGAT 1440
Db 7320 GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTTCTGGTGCCTCGGCTGAGGAA 7379

Qy	1081	ACACCAATGACAAACA	CACAGAGACA	CAATCACTG	AAATGAGCA	AAACCAAGTGC	CCCCCTCG	1140
Db	1078	ACCAACCAATGACAA	CACAGAGACA	CAATCACTG	AAATGAGCA	AAACCAAGTGC	CCCCCTCG	1137
Qy	1141	ATAACAACCCCTGCC	TCCAA	CGGGAATCCCA	CACAGCA	AAAGAGCA	CCAGCAGCA	1200
Db	1138	ATAACAACCCCTGCC	TCCAA	CGGGAATCCCA	CACAGCA	AAAGAGCA	CCAGCAGCA	1197
Qy	1201	GGCCCCGCCCA	CAACGGCA	CCAAACACAGCA	AAATGAGCATTTCC	ACAGTCTCT	CCCCCACC	1260
Db	1198	GGCCCCGCCCA	CAACGGCA	CCAAACACAGCA	AAATGAGCATTTCC	ACAGTCTCT	CCCCCACC	1257
Qy	1261	CCGAGCTCGAGTC	CACAA	CACTCTGTATATTT	TCAGAGAA	AGCGATGATTTT	CTCGAA	1320
Db	1258	CCGAGCTCGAGTC	CACAA	CACTCTGTATATTT	TCAGAGAA	AGCGATGATTTT	CTCGAGG	1317
Qy	1321	GAAGGTGATATAT	TTCCCGTTTTT	TAGATGGGT	TAAATAATCTG	AAATGATTTGAT	TGATCCA	1380
Db	1318	GAAGGCGACAT	TTCCCTTTT	CTGGATGGGT	TAAATAATGCT	CCAATGATTTG	ACCCA	1377
Qy	1381	ATCCAAACACAG	AAACAA	CAATCTTTGAT	GAATCTCCGAGCTTT	TAATCTTCA	CTAATGAG	1440
Db	1378	GTTCNAATACAA	AAACAA	CAATCTTTGAT	GAATCTCTAGT	TTCTGGTCTCGG	CTGAGGA	1437
Qy	1441	GAACAAACACT	CCCCCGAATAT	CAGATTTAACTTTCT	TATTTCTGAT	TAAATGGA	1500	
Db	1438	GATCAACAT	GCCTCCCCCA	TATTAGTTTAACTTTAT	CTTATTTCTT	TAATAATGAG	1497	
Qy	1501	GATCTGCTACT	CTGGGGAAA	ACGAGATGAT	TGTGATCAGAGTT	TGAGATTTGAGT	1560	
Db	1498	AACACTGCTACT	CTGGAGAAA	ATGAGATGAT	TGTGATCAGAGTT	TAAATTTGAGC	1557	
Qy	1561	GTGACAGAGG	ACGATTTGGG	CGAGGCTTAGCT	TGGATACCATTTT	TGGCCCTGGAATC	1620	
Db	1558	GTTCAGAGG	ATGACCTGG	CGCAGGCTCAGTT	TGGATACCGTTTTT	TGGCCCTGGAAT	1617	
Qy	1621	GAAGGACTT	ATACTGCCG	TTTTAATCA	AAATCAGAA	CAATTTAGTTTGTAG	TGTGAGG	1680
Db	1618	GAAGGACTT	ATACTGCCG	TTTTAATCA	AAATCAGAA	CAATTTAGTTTGTAG	TGTGAGG	1677
Qy	1681	CGCTTACT	TAATCAACTG	CTTAATCTT	CGAGCTCTTGT	TAGGTCAC	ACCGAGGA	1740
Db	1678	CGTCTAGC	CAATCAACTG	CTTAATCTT	TGAACTCTTAT	TGAGATGCA	CAACTGAGGA	1737
Qy	1741	AGGACATTT	CTCTTAATCA	ATAGCATG	CTATTTGCT	TACAGGTTGGG	CGGA	1800
Db	1738	AGACATTT	CTCTTAATCA	ATAGCATG	CTATTTGCT	TACCAAGATGGG	GGA	1797
Qy	1801	ACATGCAAG	TGCTAG	CACTGATTTG	CAATAGGAT	TAGAGTCTCT	CTAAAAATATC	1860
Db	1798	ACATGCAAG	TGCTAG	CACTGATTTG	CAATAGGAT	TAGAGTCTCT	CTAAAAATAT	1857
Qy	1861	TCAGACAA	ATCGACAA	AAATCAGAA	AGGATGAA	CAAAAGGAGG	AAACTGGCTGGG	1920
Db	1858	TCAGACAA	ATTTGACCA	AAATTA	AAAGGACG	AAACAA	AAAGAGGGG	1917
Qy	1921	GGTGGCAAA	TGGTGGA	CATCTG	ACTGGGTGTTCT	CAACCAATTTGGG	CACTCTGCTACTA	1980
Db	1918	GGTGGTAA	TGGTGGA	CATCCG	ACTGGGTGTTCT	CAACCAATTTGGG	CACTTTGCTACTA	1977
Qy	1981	TTATCTAT	AGCTGTTCT	GATTGCTCTG	CTTGATCTG	TGATCTG	GTGATCTTCACTAAATACAT	2040
Db	1978	TTATCCAT	AGCTGTTCT	GATTGCTCTAT	CTTGATCTT	GTGATCTTCACT	TAATATATC	2037
Qy	2041	GGATGA	2046					
Db	2038	GGATAA	2043					

RESULT 5
US-10-491-121-31
; Sequence 31, Application US/10491121

```

; Publication No. US20040259825A1
;
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
;
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
;
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
;
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
;
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
;
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 31
;
; LENGTH: 7005
;
; TYPE: DNA
;
; ORGANISM: Artificial Sequence
;
; FEATURE:
;
; OTHER INFORMATION: Synthetic Construct pVR1012x/s Marburg GP (dTM)
; US-10-491-121-31

```

Query Match	86.4%;	Score 1767.6;	DB 20;	Length 7005;	
Best Local Similarity	94.2%;	Pred. No. 0;			
Matches 1836;	Conservative	0;	Mismatches 114;	Indels 0;	Gaps 0;
Qy	1	ATGAAGACCA	CGTGTTCCTTTACGTCCTTATCTTAATTC	CAAGGGACAAAAATCTCCCC	60
Db	2034	ATGAGNCCAC	ATGTTTCTTATCAGTCCTTATCTTAATTC	CAAGGGACAAAAATCTCCCC	2093
Qy	61	ATTTTAGAGAT	AGCTAGTAAATCAACCCCAAAATGTGGATTC	CGGTATGCTCGGAAC	120
Db	2094	ATTTTAGAGAT	AGCTAGTAAATCAACCCCAAAATGTGGATTC	CGGTATGCTCGGAAC	2153
Qy	121	CTCCAGAAGAC	AGAAGCGTCCATCTGTATGGGATTCACACT	GTAGTGGGCAAAAGTTGCT	180
Db	2154	CTCCAGAAGAC	AGAAGCGTCCATCTGTATGGGATTCACACT	GTAGTGGGCAAAAGTTGCT	2213
Qy	181	GATTCCCTTTG	AGGACATCCAAAGCGATGGCTTTT	CAGGACAGGTGTACTTCCCAAGAA	240
Db	2214	GATTCCCTTTG	AGGACATCCAAAGCGATGGCTTTT	CAGGACAGGTGTACTTCCCAAGAA	2273
Qy	241	GTTGAGTACAC	AGAGGGGGAGGAGCCAAACATGCTACAATAT	AGTGTAAACGGATCCC	300
Db	2274	GTTGAGTACAC	AGAGGGGGAGGAGCCAAACATGCTACAATAT	AGTGTAAACGGATCCC	2333
Qy	301	TCTGGAAAA	TCTTGCTGTTTAGATTCCTCTTACCAACAT	CCGTGACTATCCGAAATGCAAA	360
Db	2334	TCTGGAAAA	TCTTGCTGTTTAGATTCCTCTTACCAACAT	CCGTGACTATCCGAAATGCAAA	2393
Qy	361	ACTATCCATCAT	ATATTCAGGTCAAAACCTTCATGCACAGGGAT	CGCCCTTCATTATG	420
Db	2394	ACTATCCATCAT	ATATTCAGGTCAAAACCTTCATGCACAGGGAT	CGCCCTTCATTATG	2453
Qy	421	GGAGCATTTTT	CTGTATGATCGCATTCCTCCACAAACAAT	GTATCCGAGGCAAGTCTTC	480
Db	2454	GGAGCATTTTT	CTGTATGATCGCATTCCTCCACAAACAAT	GTATCCGAGGCAAGTCTTC	2513
Qy	481	ACTGAAGGGAAC	ATAGCAGCTATGATTTGTCATTAAGACAGT	GCACAAAATGATTTCTCG	540
Db	2514	ACTGAAGGGAAC	ATAGCAGCTATGATTTGTCATTAAGACAGT	GCACAAAATGATTTCTCG	2573
Qy	541	CGGCAAGGACAA	GGGTACCGTTCATATGAATCTGACTTCTACT	AAATAATATTTGGACAAGT	600
Db	2574	CGGCAAGGACAA	GGGTACCGTTCATATGAATCTGACTTCTACT	AAATAATATTTGGACAAGT	2633
Qy	601	AGTAAACGGAA	ACGCAAAACGAATGACATCTGGATGTTT	CGGCGCTCTTTCAAGAAATACAATTCT	660
Db	2634	AGTAAACGGAA	ACGCAAAACGAATGACATCTGGATGTTT	CGGCGCTCTTTCAAGAAATACAATTCT	2693

```
QY 661 ACABAGAACCAACATGTCCTCCGTCCTCCAAATACCTCCACCACTGCCCCACAGCCGTCGG 720
Db 2694 ACNAGAACCAACATGTCCTCCGTCCTCCAAATACCTCCACCACTGCCCCACAGCCGTCGG 2753
QY 721 GAGATCAAACTCAACAGCAACCCCAACTGATGTCACCAAACTCAATPACCAAGGACCCCAAGC 780
Db 2754 GAGATCAAACTCAACAGCAACCCCAACTGATGTCACCAAACTCAATPACCAAGGACCCCAAGC 2813
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGGAGAAACGAGAAACCCCAACA 840
Db 2814 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGGAGAAACGAGAAACCCCAACA 2873
QY 841 ACTTCTGATGCGGTCCACCAAGCAAGGGCTTTCATCAACAATGCCACCCACTCCCTCACCA 900
Db 2874 ACTTCTGATGCGGTCCACCAAGCAAGGGCTTTCATCAACAATGCCACCCACTCCCTCACCA 2933
QY 901 CAACCAAGCAGCCCAACAGAGGAGAAACAACAACCACTTCCCAAGATGCTGTGACT 960
Db 2934 CAACCAAGCAGCCCAACAGAGGAGAAACAACAACCACTTCCCAAGATGCTGTGACT 2993
QY 961 GAACTAGACAAATAACACAACTGCAACCGTCCATGCGCCCTCATPAACTACCAACA 1020
Db 2994 GAACTAGACAAATAACACAACTGCAACCGTCCATGCGCCCTCATPAACTACCAACA 3053
QY 1021 ATCTCTACTTAAACACACTCCAAACCAACTTCCAGCTCTCTCTGCACTTACAAAC 1080
Db 3054 ATCTCTACTTAAACACACTCCAAACCAACTTCCAGCTCTCTCTGCACTTACAAAC 3113
QY 1081 ACCACCAATGACACACACAGAGCAACATCACTGAAATGAGCAACCAAGTCCGCCCTCG 1140
Db 3114 ACCACCAATGACACACAGAGCAACATCACTGAAATGAGCAACCAAGTCCGCCCTCG 3173
QY 1141 ATAAACAACCTGCTCCAAACGGGAAATCCCAACAGCAAGAGAGCAACAGCAAGCAAAAA 1200
Db 3174 ATAAACAACCTGCTCCAAACGGGAAATCCCAACAGCAAGAGAGCAACAGCAAGCAAAAA 3233
QY 1201 GGCCCCGCCAAGCGGCAACCAACGACCAATGAGCAATTCACCAAGTCCCTCCCCCACC 1260
Db 3234 GGCCCCGCCAAGCGGCAACCAACGACCAATGAGCAATTCACCAAGTCCCTCCCCCACC 3293
QY 1261 CCCAGCTCGACTGCAACAACATCTGTATATTTCAAGAAAGCGATCGATTTCTCGGAAA 1320
Db 3294 CCCAGCTCGACTGCAACAACATCTGTATATTTCAAGAAAGCGAGTATCTCTGGAGG 3353
QY 1321 GAAGTGATATATTCCTGTTTATAGTGGTTTAAATAATCTGAAATTTGATTTGATCCA 1380
Db 3354 GAAGGCGACATGTTCCCTTTCTGGATGGTTTAAATAATGCTCCAAATTTGATTTGACCCA 3413
QY 1381 ATCCCAACACAGAAACAACTTTTGATGAATCTCCAGCTTTAATACTCAACTAATGAG 1440
Db 3414 GTTCCAAATACAAAAACAACTTTTGATGAATCTCTAGTTCTGGTGCCCTCGGCTGAGGAA 3473
QY 1441 GAAACAACACTCCCGGCAATACAGTTTAACTTTCTCTTATTTTCTTGATAAAATGGA 1500
Db 3474 GATCAACATGCTCCCGCAATATAGTTTAACTTTATCTTATTTCTTAAATAAATGAG 3533
QY 1501 GATACTGCTACTCTGGGGAACACGAGAAATGATTTGATGACAGAGTTGAGGATTTGGAGT 1560
Db 3534 AACACTGCTACTCTGGAGAAAAATGAGAAATGATTTGATGACAGAGTTAAGAAATTTGAGC 3593
QY 1561 GTGAGAGGAGGAGATTTGGCGGAGGCTTAGCTGGATACCAATTTTGGCCCTGGGAATC 1620
Db 3594 GTTCAAGAGGAGTACCTTGCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAT 3653
QY 1621 GAAGCACTTATACCTGCGGTTTAAATCAAAATCAGAACATTTAGTTTGTAGTTGAGG 1680
Db 3654 GAAGGACCTTACCTGCTGTTTTTAAATTAATAATCAAAACAATTTGGTCTGAGGTTGAGG 3713
QY 1681 CGCTTAGCTTAATCAAACTGCTAAATCTTTGGAGCTCTTTGTTAAAGGGTCACCAACCGAGGAA 1740
Db 3714 CGTCTAGCCAATCAAACTGCGCAATCTTTGGAATCTTTATTCAGAGAGTCACAACTGAGGAA 3773
QY 1741 AGGACATTTTCTTAAATCAATAGGCAATGCAATTTGACTTTTTTGTCTTACGAGGTTGGGCGGA 1800
```

```
Db 3774 AGAACATTTCTCTTAAATCAATAGACATGCTATTGACTTTTACTCACAAAGATGGGGAGA 3833
QY 1801 ACATGCAAGGTGCTAGGACCTGATTGTTGCTAGGAATAGAAAGATCTATCTTAAAAATATC 1860
Db 3834 ACATGCAAGGTGCTAGGACCTGATTGTTGCTAGGATAGAAAGACTTTGTCCAAAAATATT 3893
QY 1861 TCAGAACAAATCGACAAATCAGAAAGATGAAACAAAGGAGGAACTGGCTGGGGTCTA 1920
Db 3894 TCAGAGCAAAATTTGACCAAAATTTAAAAAGGACGAAACAAAAAGAGGGGACTGGTTGGGGTCTG 3953
QY 1921 GGTGGCAAAATGCTGCACATCTGACTGGGGT 1950
Db 3954 GGTGGTAAATGCTGCACATCTGACTGGGGT 3983

RESULT 6
US-10-491-121-32
; Sequence 32, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 8256
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pAdapt Marburg GP (dTM)
US-10-491-121-32

Query Match 86.4%; Score 1767.6; DB 20; Length 8256;
Best Local Similarity 94.2%; Pred.No. 0;
Matches 1836; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 1 ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
Db 1431 ATGAAGACCACATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 1490
QY 61 ATTTTATAGATAGCTAGTAAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 1491 ATTTTATAGATAGCTAGTAAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 1550
QY 121 CTCAGAGAAGACAGAAAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 1551 CTCAGAGAAGACAGAAAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 1610
QY 181 GATTCCCTTTGGAGGATCCCAAGCGATGGGCTTTTCAAGACAGGTGTACCTCCCAAGAT 240
Db 1611 GATTCCCTTTGGAGGATCCCAAGCGATGGGCTTTTCAAGACAGGTGTACCTCCCAAGAT 1670
QY 241 GTTCAGTACACAGAGGGGGAGAGCCAAACATGCTACATATATAGTGTAAACGATCCC 300
Db 1671 GTTCAGTACACAGAGGGGGAGAGCCAAACATGCTACATATATAGTGTAAACGATCCC 1730
QY 301 TCTGGAAAAATCCTTGTGCTTAGATCTCTTACCACCAATCCCGTACTATCCGAAATGCAAA 360
Db 1731 TCTGGAAAAATCCTTGTGCTTAGATCTCTTACCACCAATCCCGTACTATCCGAAATGCAAA 1790
QY 361 ACTATCCATCATATATTCAGGTCAAAAACCCCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
```

[illegible]

2871	Db	GATCAACATGCCCTCCCCCAATATAGTTTAACTTTTATCTTATTTCTCTAATAATAATGAG	2939
1501	Qy	GATACTGCCTACTCTCGGGAAAAACGAAATGATTTGTATGCAGAGTTGAGGATTTGGAGT	1560
2931	Db	AACACTGCCTACTCTCGAGAAAAATGAAATGATTTGTGATGCAGAGTTAAGAAATTTGGAGC	2990
1561	Qy	GTCCAGGAGACGATTTGGCGGCAGGGCTAGCTGGATACCAATTTTTTGGCCCTCGGAATC	1620
2991	Db	GTTCAGAGAGTGAACCTTGGCCGCAGGGCTCAGTTGGATACCGTTTTTTGGCCCTCGGAATT	3050
1621	Qy	GAAGGAGCTCTATACTGCGCGTTTTAAATCAAAAAATCAGAACAAATTTAGTTTGTAGGTTGAGG	1680
3051	Db	GAAGGACTTTTACACTGCTGTTTTTAAATTAATAATCAAAAACAATTTGGTCTGCAGGTTGAGG	3110
1681	Qy	CGCTTAGCTTAATCAAACTGCTAAATCCTTGGAGCTCTGTTAAGGGTCACAACCGAGGAA	1740
3111	Db	CGTCTAGCCAAATCAAAATGCAAAATCCTTGGAACTCTTATTGAGAGTCAACAATGAGGAA	3170
1741	Qy	AGGACATTTTTCCTTAATCAATAGGCATGCAATTTGACTTTTTTGTCTACGAGTGGGGCGGA	1800
3171	Db	AGAACATTCCTCTTAATCAATAGACATGCTATTTGACTTTTCTACTCACAAGATGGGAGGA	3230
1801	Qy	ACATGCAAGGTGTAGGACCTGATTTGTCATAGGAATAGAAGATCTATCTTAAANATATC	1860
3231	Db	ACATGCAAAAGTGTGTGACCTGATTTGTGCATCGGGATAGAAGACTTGTCTCAAAAAATATT	3290
1861	Qy	TCAGAACAAATCGACAAAAATCAGAAAAGGATGAACAAAAAGGAGGAAACTGGCTGGGGTCTA	1920
3291	Db	TCAGAGCAATTTGACCAATTAATAAAGGACGAACAAAAAGAGGGGACTGGTTCGGGCTCTG	3350
1921	Qy	GGTGGCAATTTGGTGGACATCTGACTCGGGT	1950
3351	Db	GGTGGTAAATGGTGGACATCCGACTCGGGT	3380

RESULT 7

```

US-10-811-353-2
? Sequence 2, Application US/10811353
? Publication NO. US20050112098A1
? GENERAL INFORMATION:
? APPLICANT: McCray, Paul B.
? APPLICANT: Sanders, David A.
? APPLICANT: Jeffers, Scott A.
? APPLICANT: Davidson, Beverly L.
? APPLICANT: Sinn, Patrick L.
? TITLE OF INVENTION: PSEUDOTYPED VIRUSES AND METHODS FOR THEIR USE
? FILE REFERENCE: 290 00670120
? CURRENT APPLICATION NUMBER: US/10/811,353
? CURRENT FILING DATE: 2004-03-26
? NUMBER OF SEQ ID NOS: 16
? SOFTWARE: PatentIn version 3.2
? SEQ ID NO 2
? LENGTH: 2046
? TYPE: DNA
? ORGANISM: Marburg virus glycoprotein
US-10-811-353-2

```

	Query Match	82.7%	Score 1692.4	DB 21	Length 2046
	Best Local Similarity	89.2%	Pred. No. 0		
	Matches 182	Conservative 0	Mismatches 221	Indels 0	Gaps 0
Qy	1	ATGAAGACCACATGTTTCCTTATCAGTCCTTATCTTAAATCAAGGGACAAAAATCTCCCC	60		
Db	1	ATGAAGACCACATGTCCTTTTATCAGTCCTTATCTTAAATCAAGGGATAAAAATCTCTCCCT	60		
Qy	61	ATTTTAGAGATAGCTAGTAAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAACT	120		
Db	61	ATTTTAGAGATAGCTAGTAAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAACT	120		
Qy	121	CTCCAGAAGACAGAAAGCTCCATCTGATGGGATTCACTAGATGGGCAAAAAGTTGCT	180		
Db	121	CTCCAGAAGACAGAAAGATGCCATCTGATGGGATTCACTAGATGGGCAAAAAGTTGCT	180		

181 GATTCCTCTTTGGAGGATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAT 240
181 GATTCCTCTTTGGAGGATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAAT 240
241 GTTGAAGTACACAGAGGGGAGGAGCCAAACATGCTACAATAAAGTGTAAAGGATCC 300
241 GTTGAAGTACACAGAGGGGAGGAGCCAAACATGCTACAATAAAGTGTAAAGGATCC 300
301 TCTGGAATCTTGTCTGTATGATCTCTCAACATCCTGTACATATCGAATATCGAAATG 360
301 TCTGGAATCTTGTCTGTATGATCTCTCAACATCCTGTACATATCGAATATCGAAATG 360
361 ACTATCCATCATATTCAGAGTCAAAACCTCATGCGCAAGGGATCGCCCTCATTTG 420
361 ACTATCCATCATATTCAGAGTCAAAACCTCATGCGCAAGGGATCGCCCTCATTTG 420
421 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACATGTACCGAGGCAAGTCTTC 480
421 GGAGCATTTTTCTGTATGATCGCATTCCTCCCAACATGTACCGAGGCAAGTCTTC 480
481 ACTGAAGGAAACATAGCAGCTATGATGTCTCAATAGACAGTGCACAAATGATTTCTCG 540
481 ACTGAAGGAAACATAGCAGCTATGATGTCTCAATAGACAGTGCACAAATGATTTCTCG 540
541 CGGCAAGGACAGGCTACCGTCATATGATCTGACTTCTACTAATAATATGACAGT 600
541 AGGCAAGGACAGGCTACCGTCATATGATCTGACTTCTACTAATAATATGACAGT 600
601 AGTAACGGAACGCAAAATGACACTGCTGATGTTTCGGCGCTCTTCAAGAATACAAATCT 660
601 AACATGGAACCAAAACGAAATGACACTGCTGATGTTTCGGCGCTCTTCAAGAATACAAATCT 660
661 ACAAAGAACCAAAATGCTCGTCCGTCGAAATATCTCCACCACTGCCCCACAGCCGCTCG 720
661 ACAAAGAACCAAAATGCTCGTCCGTCGAAATATCTCCACCACTGCCCCACAGCCGCTCG 720
721 GAGATCAAACTCAGACAGCAACCCCACTGATGCGCAACCACTCAATACCAAGCCCAAGC 780
721 GAGATCAAACTCAGACAGCAACCCCACTGATGCGCAACCACTCAATACCAAGCCCAAGC 780
781 AGTATGATGAGGACCTCGCAACATCGCGCTCAGGTCGCGAGAACAGAACCCCAACA 840
781 AATGATGATGAGGACCTCGCAACATCGCGCTCAGGTCGCGAGAACAGAACCCCAACA 840
841 ACTTCTGATGCGGTCAACAAGCAAGGCTTTTCATCAAAATGCCCCCACTCCCTCACCA 900
841 ACTTCTGATGCGGTCAACAAGCAAGGCTTTTCATCAAAATGCCCCCACTCCCTCACCA 900
901 CAACCAAGCACGCCACAGAGGAGGAAACCAACAAACCAATTCCTCCAAAGTGTGTGACT 960
901 CAACCAAGCACGCCACAGAGGAGGAAACCAACAAACCAATTCCTCCAAAGTGTGTGACT 960
961 GAACTAGACAAATAACCAACTCAGACCGTCCATGCGCCCTCATACCACTACCA 1020
961 GAACTAGACAAATAACCAACTCAGACCGTCCATGCGCCCTCATACCACTACCA 1020
1021 ATCTCTACTTAAACAACCTCCAAACAACCTTTCAGCACTCTCTGTCACCAATTAACAAAC 1080
1021 ATCTCTACTTAAACAACCTCCAAACAACCTTTCAGCACTCTCTGTCACCAATTAACAAAC 1080
1081 ACCACCAATGACACACAGAGGACCAATCATCTGAAATGAGCAAAACAGTGTCCCTCG 1140
1081 ACCACCAATGAGACACACAGAGGACCAATCATCTGAAATGAGCAAAACAGTGTCCCTCG 1140
1141 ATAAACCCCTGCTCCCAAGGAAATCCCAACAGCAAGAGACAGCAGGAGCAAA 1200
1141 ATAAACCCCTGCTCCCAAGGAAATCCCAACAGCAAGAGACAGCAGGAGCAAA 1200
1201 GGGCCCGCCACCAACGCGCACCAACAGCAAAATGAGCAATTCACCAAGTCTCCCTCCAC 1260
1201 GGGCCCGCCACCAACGCGCACCAATATGACAAATGGGCAATTAACCAAGTCTCCCTCCAC 1260
1261 CCCAGCTCGACTGCAACAATCTTGATATATTTTCAAGAAAGCGATCGATTTCTGGA 1320

1261 CCCAGCTCGACTGCAACAATCTTGATATATTTTCAAGAAAGCGATATCTCTGAGG 1320
1321 GAAGGTGATATATTTCCGTTTTTAGATGGGTAAATAAATACTGAAATTTGATTTGATCCA 1380
1321 GAAGGTGATATATTTCCGTTTTTAGATGGGTAAATAAATACTGAAATTTGATTTGATCCA 1380
1381 ATCCCAACACAGAAACAATCTTTGATGATATCTCCGAGCTTTAATACTTCAACTAATGAG 1440
1381 GTTCCAAATCAAGACGATCTTTGATGATATCTTCTAGTTCTGGTCTTCGGCTGAGGAA 1440
1441 GAACACACACTCCCGGAAATATCAGTTTAACTTTCTCTTATTTTCTGATAAAATGGA 1500
1441 GATCAACATGCTCTCCCAATATCAGTTTAACTTTATCTTATTTCTTAATATAAATGAA 1500
1501 GATCTGCTACTCTCGGGGAAACGAGATGATTTGATGACAGAGTTTGAGGATTTGGAGT 1560
1501 AACACTGCTACTCTCGGAGAAATGAGAACGATTTGATGACAGAGTTTAAGAAATTTGGAGC 1560
1561 GTGAGGAGGACGATTTGGCGGACGGCTTAGCTGGATACATTTTGGCCCTGGAATC 1620
1561 GTTCAGGAGGATGACCTGGCAGCAGGCTCAGTTGGATACCGTTTGGCCCTGGAATC 1620
1621 GAAGACTCTATCTGCGGTTTAAATCAAAATCAGAAACAATTTAGTTGTAGTTGAGG 1680
1621 GAAGACTCTTATCTGCTGTTTAAATTAATAAACCAGAAACAATTTGGTCTGCGGTTGAGG 1680
1681 CGCTTACTTAACTGCTTAAATCTTGGAGCTCTTGTAAAGGCTCACAACCGAGGAA 1740
1681 CGCTAGCCAAATCAAACTGCAAAATCTTGGAACTCTTATTAAGAGTCACAACCGAGGAA 1740
1741 AGGACATTTCTTAAATAGAGCATGCAATTTGACTTTTGTCTTACGAGTGGGGCGGA 1800
1741 AGGACATTTCTTAAATTAATAGACATGSCCAATGACTTTCTACTACAAGGTGGGGAGGA 1800
1801 ACATGCAAGGTGCTAGGACCTGATTTGTCATAGAGATAGAGATCTATCTTAAATAATC 1860
1801 ACATGCAAGGTGCTTGGACCTGATTTGTCATAGAGATAGAGATCTTGTCCAGGAATAT 1860
1861 TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAGGAGGAAACTGGCTGGGGCTTA 1920
1861 TCAGAACAAATTCGACAAATCAGAAAGGATGAAACAAAGGAGGAACTGGCTGGGGCTTA 1920
1921 GGTGGCAATGTTGAGACATCTGACTGGGTGTTCTTCAACCAATTTGGGCATCTCTGCTACTA 1980
1921 GGTGGTAAATGTTGAGACATCTGACTGGGTGTTCTTACTAACTTGGGCATTTTGGCTACTA 1980
1981 TTATCTATAGTGTCTGATGCTCTGCTCTGCTATCTGCTGATCTTCACTAATATTAAT 2040
1981 TTATCCATAGTGTCTGATGCTCTATCTCTGCTATTTGCTGATCTTTTACCAATAATATC 2040
2041 GGATGA 2046
2041 GGGTAA 2046

RESULT 8

US-10-353-856-27

; Sequence 27, Application US/10353856

; Publication No. US20030215794A1

; GENERAL INFORMATION:

; APPLICANT: Kawaoka, Yoshihiro

; APPLICANT: Jasenosky, Luke D.

; APPLICANT: Neumann, Gabriele

; APPLICANT: Wisconsin Alumni Research Foundation

; TITLE OF INVENTION: Filovirus Vectors and No. US20030215794A1infectious Filovirus-Base

; FILE REFERENCE: 800.032US1

; CURRENT APPLICATION NUMBER: US/10/353,856

; PRIOR FILING DATE: 2003-01-29

; PRIOR APPLICATION NUMBER: US 60/353,972

; PRIOR FILING DATE: 2002-01-31

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 27									
; LENGTH: 19112									
; TYPE: DNA									
; ORGANISM: Marburg virus									
US-10-353-856-27									
Query Match 82.7%; Score 1692.4; DB 17; Length 19112;									
Best Local Similarity 89.2%; Pred. No. 0;									
Matches 1825; Conservative 0; Mismatches 221; Indels 0; Gaps 0;									
QY	1	ATGAAGACACATGTTCTTATCAGTCTTATCTTAAATTCAGGACAAAAATCTCCC	60						
DB	5940	ATGAAGACACATGTTCTTTATCAGTCTTATCTTAAATTCAGGACAAAAATCTCCC	5999						
QY	61	ATTTTATGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	120						
DB	6000	ATTTTATGAGATAGCTAGTAACATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	6059						
QY	121	CTCCAGAGACAGAGAGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT	180						
DB	6060	CTCCAGAGACAGAGAGATGCTCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT	6119						
QY	181	GATTCCTCTTTGGAGCATCCAGCGATGGCTTTTCCAGACAGGTGTACCTCCCAAGAT	240						
DB	6120	GATTCCTCTTTGGAGCATCCAGCGATGGCTTTTCCAGACAGGTGTACCTCCCAAGAT	6179						
QY	241	GTGTAGTACACAGAGGGGAGGAGCCAAAACATGCTACAATATAAGTGTAAACGGATCCC	300						
DB	6180	GTGTAGTATACAGAGGGGAGGAGCCAAAACATGCTACAATATAAGTGTAAACGGATCCC	6239						
QY	301	TCGTGAAATCTTGTGTATGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA	360						
DB	6240	TCGTGAAATCTTGTGTATGATCCTCTACCAACATCCGTGACTATCCGAAATGCAAA	6299						
QY	361	ACTATCCATCATATTCNAGGTCAAAACCTCATGACACAGGGGATCGCCCTTCATTTAG	420						
DB	6300	ACTATCCATCATATTCNAGGTCAAAACCTCATGACACAGGGGATCGCCCTTCATTTAG	6359						
QY	421	GGAGCATTTTTCTGTATGATCGCATTCCTCCACAAATGTACCGAGGCAAGTCTTC	480						
DB	6360	GGAGCATTTTTCTGTATGATCGCATTCCTCCACAAATGTACCGAGGCAAGTCTTC	6419						
QY	481	ACTGAAGGAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTTCTCG	540						
DB	6420	ACTGAAGGAACATAGCAGCTATGATGTCAATAAGACAGTGCACAAAATGATTTCTCG	6479						
QY	541	CGGCAAGGACAGGGTACCGTCAATATGATCTGACTTCTACTAATAATATGGCAAGT	600						
DB	6480	AGGCAAGGACAGGGGTACCGTCAATATGATCTGACTTCTACTAATAATATGGCAAGT	6539						
QY	601	AGTAACGGAAACGAAACGAAATGACACTGGATGTTTCGGCGCTCTTCAAGAATACAAATCT	660						
DB	6540	AACATGGNACACAAACGAAATGACACTGGATGTTTCGGTGCTCTTCAAGNATACAACTCC	6599						
QY	661	ACAAAGAACCAACATGTGCTCCGTCCAAAATACCTCCACCACTGCCCAACGCCGCTCG	720						
DB	6600	ACGAAGAATCAAAACATGTGCTCCGTCCAAAATACCTCCACCACTGCCCAACGCCGCTCA	6659						
QY	721	GAGATCAACTCACAAGCACCCCACTGATGCCACCACTCAATACCAAGCCCAAGC	780						
DB	6660	GAGATCAAAACCAAGCACCCCACTGATGCCACCACTCAATACCAAGCCCAAGC	6719						
QY	781	AGTGTATGAGGACCTCGCAACATCCGGCTCAGGGTCCGAGAGACGAGAACCCCAACA	840						
DB	6720	ATGTATGATGAGACCTCATACATCCGGTTCAGGGTCCGAGAGACGAGAACCCCTATACA	6779						
QY	841	ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTCATCAACAATGCCACCACTCCCTCACA	900						
DB	6780	ACTTCTGATGCGGTCACTAAGCAAGGGCTTTCATCAACAATGCCACCACTCCCTCACA	6839						
QY	901	CAACCAAGCACGCCACAGCAAGGAGGAAACACACAAACCATTTCCCAAGATGCTGTGACT	960						
DB	6840	CAACCAAGCACGCCACAGCAAGGAGGAAACACACAAACCATTTCCCAAGATGCTGTGACT	6899						
QY	961	GAACCTAGACAAAAATAACAACTGCAACAACCGTCCATGCCCCCTCATAAACACTACCAACA	1020						
DB	6900	GAACCTAGACAAAAATAACAACTGCAACAACCGTCCATGCCCCCTCATAAACACTACCAACA	6959						
QY	1021	ATCTCTACTAACAACACCTCCAAACACAACTTTCAGCACTCTCTCTGCAACCAATACAAAC	1080						
DB	6960	ATCTCTACTAACAACACCTCCAAACACAACTTTCAGCACTCTCTCTGTAATCACTACAAAC	7019						
QY	1081	ACCACCAATGACACACACAGAGACCAATCACTGAAATAGCAAAACAGTCCGCCCTCG	1140						
DB	7020	ACCACCAATGACACACACAGAGACCACTGAAATAGCAAAACAGTCCGCCCTCG	7079						
QY	1141	ATAACAACCTCTCCAAACCGGAAATCCCAACAGCAAAAGAGCACAGCAGCAAAAAA	1200						
DB	7080	ATAACAACCTCTCCAAACAGGAAATCTTACCACAGCAAAAGAGCACTTAAACAACGAAA	7139						
QY	1201	GGCCCCGCCACAAACCGGACCAACACAGCAAAATGAGCAATTCACAGTCCCTCCGCCACC	1260						
DB	7140	GGCCCCGCCACAAACCGGACCAAAATATGCAAAATGGCATTTAACAGTCCCTCCGCCACC	7199						
QY	1261	CCCAGCTCGACTGCACAACTCTTTGTATATTTCAAGAAAGCGATCGATTTCTCGAAA	1320						
DB	7200	CCCAGCTCGACTGCACAACTCTTTGTATATTTCAAGAAAGCAAGATATCTCTCGAGG	7259						
QY	1321	GAAGGTGATATATTTCCGTTTTTAGATGGGTTAATAAATCTGAAATTTGATTTGATCCA	1380						
DB	7260	GAAGGTGATATATTTCCGTTTTTAGATGGGTTAATAAATCTGAAATTTGATTTGATCCA	7319						
QY	1381	ATCCAAACACAGAAACATCTTTGATGAAATCTCCAGCTTTAATCTCAACTAATGAG	1440						
DB	7320	ATCCAAACACAGAAACATCTTTGATGAAATCTCCAGCTTTAATCTCAACTAATGAG	7379						
QY	1441	GAAACAACACTCCCGCAATATCAGTTTAACTTTCTTATTTTCTGATAAAATGGA	1500						
DB	7380	GATCAACATGCTCCCAATATCAGTTTAACTTTTCTTATTTCTGATAAAATGGA	7439						
QY	1501	GATCTGCTACTCTGGGGAACCGAGAAATGATTTGATGAGAGTTTGAGGATTTGGAGT	1560						
DB	7440	AACACTGCTACTCTGGGGAACCGAGAAATGAGAACGATTTGATGAGAGTTTGGAGC	7499						
QY	1561	GTGAGAGGAGAGATTTGGCGGAGGGCTTAGCTGGATACCATTTTGGCCCTGGAATC	1620						
DB	7500	GTTCAGGAGGATGACCTGGCAGAGGGCTCAGTTGGATACCGTTTTTGGCCCTGGAATC	7559						
QY	1621	GAAGACTCTATATCTGCGGTTTTAATCAAAATCAGAAACATTTAGTTTGTAGGTTGAGG	1680						
DB	7560	GAAGACTTTTATCTGCTGGTTTTAATTAACCAAAACAAATTTGGTCTGCAGGTTGAGG	7619						
QY	1681	CGCTTAGCTAATCAAACTGCTAAATCTCTGGAGCTCTTTGTTAAGGGTCAACACGAGAA	1740						
DB	7620	CGCTTAGCTAATCAAACTGCTAAATCTCTGGAGCTCTTTGTTAAGGGTCAACACGAGAA	7679						
QY	1741	AGGACATTTCTTAACTAATAGCATGCAATTCACCTTTTGGCTTACGAGTGGGGCGGA	1800						
DB	7680	AGGACATTTCTTAACTAATAGCATGCAATTCACCTTTTGGCTTACGAGTGGGGCGGA	7739						
QY	1801	ACATGCAAGGTGCTAGACACCTGATTTGTGATAGAAATAGAGATCTATCTTAAATAATC	1860						
DB	7740	ACATGCAAGGTGCTAGACACCTGATTTGTGATAGAAATAGAGATCTTGTCCAGGNATAT	7799						
QY	1861	TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAGGAGGAAACTGGCTGGGGCTTA	1920						
DB	7800	TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAGGAGGAACTGGTGGGGCTTA	7859						
QY	1921	GGTGGCAATGGTGGACATCTGACTGGGTGTTCTCACCATTTGGGGATCTCTGCTACTA	1980						
DB	7860	GGTGGTAAATGGTGGACATCTGACTGGGTGTTCTTACTAACTTTGGGATTTGCTACTA	7919						
QY	1981	TTATCTATAGTGTCTTGATTTGCTCTGTCGTGATCTCTGCTATCTTCACTAAATACAT	2040						
DB	7920	TTATCTATAGTGTCTTGATTTGCTCTATCCGTATTTGCTGTAATTTTACCAATATATC	7979						

Qy 2041 GGATGA 2046
Db 7980 GGGTAA 7985

RESULT 9

US-10-066-506A-13
; Sequence 13, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan. L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 13
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Marburg virus strain Raven

; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 and Marburg
; OTHER INFORMATION: Glycoprotein 2
US-10-066-506A-13

Query Match 74.2%; Score 1518; DB 15; Length 2046;
Best Local Similarity 83.9%; Pred. No. 0;
Matches 1716; Conservative 0; Mismatches 330; Indels 0; Gaps 0;

Qy 1 ATGAAGACACATGTTTCCTTATAGTCTTATCTTAATTCAGGACAAAAAATCTCCCC 60
Db 1 ATGAAGACCATATATTTCTGATAGTCTCATTTTATCCAAAGTATAAAAACTCTCCCT 60
Qy 61 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Db 61 GTTTTAGAAATTTGCTAGTAACAGCCAACTCAAGATGTAGATTCAAGTGTGCTCCGGAAC 120
Qy 121 CTCGAGAGACAGAGACGTTCATCTGATGGATTTCACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCGAGAGACAGAGAGTTCATCTGATGGATTTTACATGAGTGGGCAAAAAGTTGCT 180
Qy 181 GATTCCTCTTTCGAGGSCATCCAAAGCATGGGCTTTTCAGACAGGTGTACCTCCCAAGAT 240
Db 181 GATTCCTCTTTCGAGAGCATCTAAGCATGGGCTTTTCAGACAGGTGTCTCTCCCAAGAAC 240
Qy 241 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACAATATAAGTGTAAACGGATCCC 300
Db 241 GTTGAGTATACGAGGAGAGAGCCAAACATGTTACAATATAAGTGTAAACAGCCTC 300
Qy 301 TCTGGAATCTTCTGCTGTGTAGATCTCTCAACATCCGTTGATCTCCGAAATGCAAA 360
Db 301 TCTGGAATCTTCTGCTGTGTATCTCCAGTAATATCCGGAATATACCTTAATGTAA 360
Qy 361 ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTGTTCATCATATTCAGGTCAAAACCTCATGACAGGGGATTCGCTCCATTTGTGG 420
Qy 421 CGAGCATTTTTCTGTATGATCGATTCCTCAACAAATGTACCGAGGCAAGTCTTC 480
Db 421 GGGGCATTTTTCTGTATGATCGGTGCTCTACAACTATGACCGAGGCAAGTCTTC 480
Qy 481 ACTGAAGGAACATAGCAGCTATGATGTCAATAGACAGTGCACAAAATGATTTCTCG 540
Db 481 ACTGAAGGAATATAGCAGCTATGATGTCAATAGACAGTTCACAGATGATTTTTCT 540
Qy 541 CGGCAAGGACAGGGTACCGTATATGATCTGACTTCTACTAATAAATATTTGGACAAGT 600
Db 541 AGGCAAGGACAGGGTATATCGTCACATGAACCTTGACCTCCACCANTAAATATTTGGACAAGC 600

Qy 601 AGTAAACGGAACGCAACGAATGACACATGGATGTTTTCGGCGCTCTTCAAGAATACAAATCT 660
Db 601 AGCAATGAACGCGAGAGAAATGATACGGGATGTTTTCGGCATCTCTCCRAAGAATACAACTCC 660
Qy 661 ACAAGAACCAAAACATGTGCTCCGTCGCAAAATACCTCCACCACATGCCCCACAGCCGCTCG 720
Db 661 ACAAAATCAAAATCAATGCCCTCCATCTCTTAACTCCATCCCTGCCCCACAGTAATCTCG 720
Qy 721 GAGATCAAACTCACAAGCACCCCAACTGATGCCCAAACTCAATACACAGGACCAAGC 780
Db 721 AGCATTCACCTACAAATACCTCAAAATTAATAGCTAAATCTGGAATCTATGAACCAAGT 780
Qy 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGAGAACGAGAACCCCAACACA 840
Db 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Qy 841 ACTTCTGATGGGTCCACCAAGCAAGGCTTTCATCAACAAATGCCACCACCTCCCTCACA 900
Db 841 ACTCTTAATGTAGTCACTGAACAGAAACATGCTCAACAAATATGTTCCACTCTCTCACTA 900
Qy 901 CAACCAAGCACGCCACAGCAAGGAGGAAACAAACAAACCAATTCACCAAGATGCTGTGACT 960
Db 901 CATCCAGCACCTCACAACATGAGCAAAACAGTAGAATCTCTCCGACATGCTGTAATCT 960
Qy 961 GAACTAGACAAAATTAACAACTGCGACAAACCGTCCATGCCCTCCCTCATACCACTACCACA 1020
Db 961 GAGCACAAATGGAACCGACCCCAACCAACACACAGCAACGCTCTCTCAACAAATATTAATACA 1020
Qy 1021 ATCTCTACTRAACAACCTCCAAAACACAACTTCAGCACTCTCTGCGACCACTTACAAAAC 1080
Db 1021 ACTCCACCTATAACACTCTCAAGTACAACTCTAGTACTCTCTCCCTCCCAACCGCAAC 1080
Qy 1081 ACCACCAATGACAAACACAGAGACCAATCACTGAAAATGAGCAAAACAGATGCCCTCTCG 1140
Db 1081 ATCCCAATATGATACACAGCTGAACTAGCAGAAACGCAACAAACCAATGCTCAGTTG 1140
Qy 1141 ATAAACAACCTGCTCCAAACGGGAAATCCCAACAGCAAAAGAGACACAGAGGCAAAAAA 1200
Db 1141 AACCAACTCTAGATCCCAACAGAAAATCCCAACAGCAAGACACCAACAGCAACAACC 1200
Qy 1201 GGCCCCGACACAGGACCAACACAGCAAAATGAGCAATTTCCAGCTCTCTCCCTCCACC 1260
Db 1201 AACATCATGACGACATCAGATATACAAACGCAAAACCCCAACAAATTTCTCTCCGAT 1260
Qy 1261 CCCAGCTCGACTGCAACAACTCTTGTAATTTTCAGAAAGAAAGCGATCGATTTCTTGGAAA 1320
Db 1261 TCTAGTCCGACAAACCCGCTCTCTATATCTTTAGAAAGAAACGATCGATTTCTTGGAAA 1320
Qy 1321 GAAAGTGATATATTCCTCGTTTATAGATGGTTAATAAATACGTAAATTTGATTTGATCCA 1380
Db 1321 GAAAGTGATATATTCCTCGTTTATAGATGGTTAATAAATACGTAAATTTGATTTGATCCA 1380
Qy 1381 ATCCCAACACAGAAACAACTCTTTGATGAATCTCCAGCTTTAACTTTCACTCACTAATGAG 1440
Db 1381 ATCCCAACACAGAAACAACTCTTTGATGAATCTCCAGCTTTAACTTTCACTCACTAATGAG 1440
Qy 1441 GAAACAACACTCTCCCGCAATATCAAGTTTAACTTTCTTATTTCTGATGATAAAATGGA 1500
Db 1441 GAAACAACACTCTCCCGCAATATCAAGTTTAACTTTCTTATTTCTGATGATAAAATGGA 1500
Qy 1501 GATCTGCTACTCTGCGGAAAAACGAGAAATGATGTGATGACAGATTTGAGGATTTGGAGT 1560
Db 1501 GATCTGCTACTCTGCGGAAAAACGAGAAATGATGTGATGACAGATTTGAGGATTTGGAGT 1560
Qy 1561 GTGAGGAGGAGGATTTGCGCGCAGGCTTAGCTGGATACCAATTTTGGCCCTCGGAATC 1620
Db 1561 GTGAGGAGGAGGATTTGCGCGCAGGCTTAGCTGGATACCAATTTTGGCCCTCGGAATC 1620
Qy 1621 GAAGGACTCTTACTGCGCGTTTAAATCAAAAATCAGAAACAAATTTAGTTTGTAGGTTGAGG 1680
Db 1621 GAAGGACTCTTACTGCGCGTTTAAATCAAAAATCAGAAACAAATTTAGTTTGTAGGTTGAGG 1680

```
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCCCTTGGAGCTCTTGTAAAGGGTCACAACCGAGAA 1740
Db |||||
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCCCTTGGAGCTCTTGTAAAGGGTCACAACCGAGAA 1740
Db |||||
QY 1741 AGGACATTTTCCTTAATCAATAGGCGATGCAATGACTTTTTCCTTACGAGGTGGGGCGGA 1800
Db |||||
QY 1741 AGGACATTTTCCTTAATCAATAGGCGATGCAATGACTTTTTCCTTACGAGGTGGGGCGGA 1800
Db |||||
QY 1801 ACATGCAAGGTGCTAGGACCTGATTGTTCATAGGAATAGAAAGATCTATCTAAAAATATC 1860
Db |||||
QY 1801 ACATGCAAGGTGCTAGGACCTGATTGTTCATAGGAATAGAAAGATCTATCTAAAAATATC 1860
Db |||||
QY 1861 TCAGAAACAAATCGACAAATAGAAAGGATGAAACAAAAGGAGAAACTGGCTGGGGTCTA 1920
Db |||||
QY 1861 TCAGAAACAAATCGACAAATAGAAAGGATGAAACAAAAGGAGAAACTGGCTGGGGTCTA 1920
Db |||||
QY 1921 GGTGGCAATGCTGGACATCTGACTGGGGTGTCTCACCMAATTTGGGCATCTGCTACTA 1980
Db |||||
QY 1921 GGTGGCAATGCTGGACATCTGACTGGGGTGTCTCACCMAATTTGGGCATCTGCTACTA 1980
Db |||||
QY 1981 TTATCTATAGCTGTTCTGATTGCTCTGCTCTGTATCTGTGCTATCTTCACTAAATACATT 2040
Db |||||
QY 1981 TTATCTATAGCTGTTCTGATTGCTCTGCTCTGTATCTGTGCTATCTTCACTAAATACATT 2040
Db |||||
QY 2041 GGATGA 2046
Db |||||
QY 2041 GGATGA 2046
Db |||||
```

RESULT 10

```
US-10-066-506A-7
; Sequence 7, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 7
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Raven Glycoprotein
; OTHER INFORMATION: Marburg virus strain Musoke Glycoprotein 2
US-10-066-506A-7
```

```
Query Match 64.5%; Score 1319.6; DB 15; Length 2046;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1592; Conservative 0; Mismatches 454; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTTCCTTATCAGTCTTATCTTAAATTCAGGGGACAAAAATCTCCCC 60
Db |||||
QY 1 ATGAAGACCAATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAAACTCTCCCT 60
Db |||||
QY 61 ATTTTAGATAGTAGTAAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAACT 120
Db |||||
QY 61 GTTTTAGAAATTTGCTAGTAACGCCAACCTCAAGATGATGATTCAGTGTGCTCCGGAACT 120
Db |||||
QY 121 CTCGAGACAGAGAGACGCTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 180
Db |||||
QY 121 CTCGAGACAGAGAGATGCTCATCTGTGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db |||||
QY 181 GATTCCTCTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db |||||
QY 181 GATTCCTCTTTGGAGGCATCTAAACGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAT 240
Db |||||
```

```
QY 241 GTTGAGTACACAGAGGGGAGGAAGCCAAACATGCTACAAATATAAGTGTAAACGGATCCC 300
Db |||||
QY 241 GTTGAGTATACGGAAGGAGGAAGCCAAACATGTTACAATATAAGTGTAAACGACCT 300
Db |||||
QY 301 TCTGAAAAATCCTTGTCTGTAGATCCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Db |||||
QY 301 TCTGAAAAATCCTTGTCTGTAGATCCCTCTACCAACATCCGCTGACTATCCGAAATGCAAA 360
Db |||||
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db |||||
QY 361 ACTGTTCAATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATTCGCTTCATTTGTTG 420
Db |||||
QY 421 GGAGCATTTTTCTGTATGATCGCATTTGCCCTCCACAAATGTATCCGAGGCAAAAGCTTTC 480
Db |||||
QY 421 GGAGCATTTTTCTGTATGATCGCATTTGCCCTCCACAAATGTATCCGAGGCAAAAGCTTTC 480
Db |||||
QY 481 ACTGAAGGGAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCTCG 540
Db |||||
QY 481 ACTGAAGGGAATATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCTCG 540
Db |||||
QY 541 CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTATAAATATTTGGACAAGT 600
Db |||||
QY 541 AGGCAAGGACAAAGGTATTCGTCATGAACTTTGACCTCCACCAATATAATTTGGACAAGC 600
Db |||||
QY 601 AGTAAACGGAACGCAAAACGAATGACATGCGATGTTTTCGGCGCTCTTCAAGAATACAAATTC 660
Db |||||
QY 601 AGCAATGAACGACAGAGAAATGATACGGGATGTTTTCGGCATCTCCCAAGAATACAACTCC 660
Db |||||
QY 661 ACAAAGAACCAAAACATGCTCGTCCAAATATCTCCACCATGCCCCACAGCCGCTCG 720
Db |||||
QY 661 ACAAACCAATCAAAACATGCTCGTCCAAATATCTCCACCATGCCCCACAGTAATCTCG 720
Db |||||
QY 721 GAGATCAAACTCACAAGCACCCCAACTGATGCGCAAACTCAATACCAAGCCCAAGC 780
Db |||||
QY 721 AGCATTCACCTTCAAAATACCTCAAAATTAATCTGCTAAATCTGGAACATGTAACCAAGT 780
Db |||||
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 840
Db |||||
QY 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db |||||
QY 841 ACTTCTGATGCGGTACCAAGCAAGGGCTTTTCATCAACAATGCCACCCACTCCCTCACCA 900
Db |||||
QY 841 ACTCTTAATGTAGTCACTGAAACAGAAAACAATCGTCAACAATATTTGCTCCACTCTTCACTA 900
Db |||||
QY 901 CAACCAAGCAGCCACAGAGGAGGAAACAAACCAACCAATTCCTCAAGATGCTGTGACT 960
Db |||||
QY 901 CATCCAAGCACCTCACAACTGAGCAAAACAGTAGCAATCTTCCGACATGCTGTAACT 960
Db |||||
QY 961 GAACTAGACAAAAATAACAACTGCAACACCGTCCATGCCCTCATGCCCTCATAAACACTACCA 1020
Db |||||
QY 961 GAGCAATGGAACCGACCCCAACAAACAAACAGCAACGCTCTCTCAACAATACTAATAACA 1020
Db |||||
QY 1021 ATCTCTACTAACACACCTCCAAACACAACTTACAGTCTCTCTGTGACCACTTACAAAC 1080
Db |||||
QY 1021 ACTCCACCTATAACACTCTCAAGTACAACTCTAGTACTCTCTTCCCTCCAAACCGCAAC 1080
Db |||||
QY 1081 ACCACCATGACACACAGAGACACATCACTGAAATGAGCAACCAAGTCCGCTCCCTCG 1140
Db |||||
QY 1081 ATCAACCAATATGATACACACGCTGAACCTAGAGAAACGCAACCAATGCTCAGTTG 1140
Db |||||
QY 1141 ATAACCAACCTCGCTCCAAACGGGAAATCCCAACCAAGCAAGAGACCAAGCAAGCAAAAAA 1200
Db |||||
QY 1141 AACCAACTCTAGATCCAAACAGAAAATCCCAACCAAGCAAGACCAACCAAGCAACACC 1200
Db |||||
QY 1201 GGCCCCGCCCAACAGGCAACCAACAGCAAAATGAGCATTTTCCAGTCTCTCCGCCAC 1260
Db |||||
QY 1201 AACATCATCATGACATCAGATATACAAAGCAACACCCCAAAATTTCTTCCGGAT 1260
Db |||||
QY 1261 CCCAGCTCGATGACACACATCTTGTATATTTTCAAGAAAGAGCGATCGATTTCTTGGAAA 1320
Db |||||
QY 1261 TCTAGTCCGACAAACCCGCCCTCTATATCTTTAGAAAGAACGATCGATCTCTTGGAGG 1320
Db |||||
```

```
QY 1321 GAAGGTGATATATCCCGTTTTAGATGGTTAATAAATACAGAAATGATTTTGATCCA 1380
Db 1321 GAAGGCGACATGTTCCCTTTCTGATGGGTTAATAAATGCTCCAAATGATTTTGACCCA 1380
QY 1381 ATCCAAACACAGAAACAATCTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG 1440
Db 1381 GTTCCAAATACAAAACAATCTTTGATGAATCTCTAGTTCTGGTGCCCTCGGCTGAGAA 1440
QY 1441 GAAACAACACATCCCGCCGAATACAGTTTAACTTTCTTATTTTCTTGATGATAAATAATGA 1500
Db 1441 GATCAACATGCTCCCGCCGAATACAGTTTAACTTTCTTATTTTCTTGATGATAAATAATGA 1500
QY 1501 GATCTGCTACTCTGCGGGAACACAGAAATGATTTGATGATGAGAGTTCAGAGATTTGGAGT 1560
Db 1501 AACCTGCTACTCTGAGGAATAAGAAATGATTTGATGATGAGAGTTCAGAGATTTGGAGC 1560
QY 1561 GTGAGGAGGAGGATTTGCGCGCAGGCTTAGCTGGATACCAATTTTTTGGCCCTGGAAATC 1620
Db 1561 GTTCAGGAGGATGACCTGCGCGCAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAT 1620
QY 1621 GAAGGACTTATACACTGCTGTTTAAATCAAAAATCAGAAATCAAAATCAAAATCAAAATCA 1680
Db 1621 GAAGGACTTATACACTGCTGTTTAAATCAAAAATCAGAAATCAAAATCAAAATCAAAATCA 1680
QY 1681 CGCTTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTTGTTAAGGCTCACAAACGAGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTTGTTAAGGCTCACAAACGAGAA 1740
QY 1741 AGGACATTTTCTTAATCAATAGGATGCAATTTGACTTTTTTACGAGTGGGGCGGA 1800
Db 1741 AGACATTTTCTTAATCAATAGATGCTATTGACTTTTACTCACAAGATGGGGAGGA 1800
QY 1801 ACATGCAAGGTGCTAGGACCTGATTTGTTGCATAGAAATAGAAATCTATCTAAAAATATC 1860
Db 1801 ACATGCAAGGTGCTAGGACCTGATTTGTTGCATAGAAATAGAAATCTATCTAAAAATAT 1860
QY 1861 TCAGAACAAATCGACAAATCAGAAAGCTGAAAGCTGAAAGGAGGAACTGGCTGGGGCTCA 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGCTGAAAGCTGAAAGGAGGAACTGGCTGGGGCTCA 1920
QY 1921 GTGGCAATGGTGACATCTGACTGGGGTGTCTCACCATAATTTGGGCATCTGCTACTA 1980
Db 1921 GTGGTAAATGGTGACATCCGACTGGGGTGTCTTACTTAATTTGGGCAATTTGCTACTA 1980
QY 1981 TTATCTATAGCTGTTCTGATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 1981 TTATCCATAGCTGCTGATTTGCTCTATCTCTGATTTGCTGCTGCTGCTGCTGCTGCTGCT 2040
QY 2041 GGATGA 2046
Db 2041 GGATAA 2046
```

RESULT 11

```
US-10-066-506a-3
; Sequence 3, Application US/10066506a
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 3
; LENGTH: 1841
; TYPE: DNA
; ORGANISM: Artificial Sequence
```

```
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein
; OTHER INFORMATION: Ebola virus Zaire Mayinga strain Glycoprotein 2
US-10-066-506a-3
```

```
Query Match          63.9%; Score 1308; DB 15; Length 1841;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
Db 1 ATGAAGACCAATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
QY 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGGAATTCGGTATGCTCCGGAAT 120
Db 61 ATTTTAGAGATAGTAGTAATAATCAACCCCAAAATGGAATTCGGTATGCTCCGGAAT 120
QY 121 CTCAGAAAGACAGAAAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCAGAAAGACAGAAAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCCAAGAT 240
Db 181 GATTCCCTTTTGGAGGCATCCAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCCAAGAT 240
QY 241 GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATNTAAGTGTAAAGGATCCC 300
Db 241 GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATNTAAGTGTAAAGGATCCC 300
QY 301 TCTCGAAAAATCTCTTGTGTAGATCCTCTACCAACATCCGCTACTATCCGAAATGCAAA 360
Db 301 TCTCGAAAAATCTCTTGTGTAGATCCTCTACCAACATCCGCTACTATCCGAAATGCAAA 360
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
Db 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATCGCCCTTCATTTATGG 420
QY 421 GGAGCATTTTTCTGTATGATCGCATTTGCCCTCCACAACATGTACCGGAGCAAGTCTTC 480
Db 421 GGAGCATTTTTCTGTATGATCGCATTTGCCCTCCACAACATGTACCGGAGCAAGTCTTC 480
QY 481 ACTCAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCTCG 540
Db 481 ACTCAAGGAAACATAGCAGCTATGATTTGTCATTAAGACAGTGCACAAATGATTTTCTCG 540
QY 541 CGGCAAGGACAAAGGATACCGTCTATGAAATCTGACTTCTACTAAATAATTTGGACAAGT 600
Db 541 CGGCAAGGACAAAGGATACCGTCTATGAAATCTGACTTCTACTAAATAATTTGGACAAGT 600
QY 601 AGTAAACGGAACGCAAAACGAATGACATGGAATGTTTGGCGCTCTTCAAGAAATACAATTCT 660
Db 601 AGTAAACGGAACGCAAAACGAATGACATGGAATGTTTGGCGCTCTTCAAGAAATACAATTCT 660
QY 661 ACAAGAACCAAAACATGTTCTCCGTCCAAATAATCTCCACCATCGCCACACGCGCTCG 720
Db 661 ACAAGAACCAAAACATGTTCTCCGTCCAAATAATCTCCACCATCGCCACACGCGCTCG 720
QY 721 GAGATCAAACTCACAGCAACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780
Db 721 GAGATCAAACTCACAGCAACCCCAACTGATGCCACCAAACTCAATACCAAGGACCAAGC 780
QY 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
Db 781 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCCAACA 840
QY 841 ACTTCTGATGCGGTCCACNAGCAGGGGCTTTCATCAAAATGCCACCTCCCTCACA 900
Db 841 ACTTCTGATGCGGTCCACNAGCAGGGGCTTTCATCAAAATGCCACCTCCCTCACA 900
QY 901 CAACCAAGCACGCCACAGCAGAGGAGGAAACAAACAAACCAATTCAGAGATGCTGTGACT 960
Db 901 CAACCAAGCACGCCACAGCAGAGGAGGAAACAAACAAACCAATTCAGAGATGCTGTGACT 960
```

QY 961 GAACTAGACAAAATAACAACTGCAAAACCGTCCATGCCCCCTCATAAACAATAACCA 1020
DB 961 GAACTAGACAAAATAACAACTGCAAAACCGTCCATGCCCCCTCATAAACAATAACCA 1020
QY 1021 ATCTCTACTTAACAACACCTTCCAAACACAACTTTCAGCACTCTCTGCAACCAATTA 1080
DB 1021 ATCTCTACTTAACAACACCTTCCAAACACAACTTTCAGCACTCTCTGCAACCAATTA 1080
QY 1081 ACCACCAATGACAAACACACAGAGCAAACTCACTGAAATGAGCAAAACAGTCCCTCG 1140
DB 1081 ACCACCAATGACAAACACACAGAGCAAACTCACTGAAATGAGCAAAACAGTCCCTCG 1140
QY 1141 ATAAACAACCTGCTCCAAACGGGAAATCCACCAAGAGAGAGAGAGAGAGAGAGAG 1200
DB 1141 ATAAACAACCTGCTCCAAACGGGAAATCCACCAAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 GGGCCGGCCACAGCGGCAACAGGCAAACTGATGATGATGATGATGATGATGATGAT 1260
DB 1201 GGGCCGGCCACAGCGGCAACAGGCAAACTGATGATGATGATGATGATGATGATGAT 1260
QY 1261 CCAGCTGACTGCAACAACTCTGTATATTTTCAAGAGAGAGAGAGAGAGAGAGAG 1308
DB 1261 CCAGCTGACTGCAACAACTCTGTATATTTTCAAGAGAGAGAGAGAGAGAGAGAG 1308

RESULT 12

US-10-491-121-42
; Sequence 42, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; PRIOR FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 6902
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012x/s Marburg (codon optimized)
US-10-491-121-42

Query Match 48.7%; Score 995.6; DB 20; Length 6902;

Best Local Similarity 69.5%; Pred. No. 1.1e-284;

Matches 1352; Conservative 0; Mismatches 594; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATGTTTCTTATCATGCTTATTAATTAATTAATTAATTAATTAATTAATTA 60
DB 1923 ATGAAGACCAATGTTTCTTATCATGCTTATTAATTAATTAATTAATTAATTAATTA 1982
QY 61 ATTTTAGAGATAGCTAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 120
DB 1983 ATCTTGGAGATGCCAGAC 2042
QY 121 CTCGAGAGACAGAGACGCTCATCTGATGGGATTCACATGATGGGAGAGAGAGAGAGAG 180
DB 2043 CTCGAGAGACAGAGACGCTCATCTGATGGGATTCACCTGATGGGATTCACCTGATGG 2102
QY 181 GATTCCTCTTCGAGGATCCAGGATGGGATTCAGGAGAGAGAGAGAGAGAGAGAGAGAT 240
DB 2103 GACAGCCCTCTGGAGGCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2162

QY 241 GTTGAGTACACAGAGGGGGAG 300
DB 2163 GTTGAGTACACAGAGGGGGAG 2222
QY 301 TCTGGAAATCCTTGTCTGTAGATCCCTTACCAACATCCCTGACTATCCGAAATGCAAA 360
DB 2223 AGCGCAAG 2282
QY 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACAGAGGGGATCGCCCTTCATTTAT 420
DB 2283 ACCATCCACCATCTCCAGGGCCAGAACCTCATGCGCCAGGGGATCGCCCTTCACCTG 2342
QY 421 GGAGCATTTTTCTGTATGATGCTGCTCCCAACAATGTATGAGAGAGAGAGAGAGAG 480
DB 2343 GGGCGCTCTTCTGTATGATGCTGCTCCCAACAATGTATGAGAGAGAGAGAGAGAGAGAG 2402
QY 481 ACTGAAGGAGACATAGAGCTATGATGCTCAATGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 2403 ACCGAGGCAACATCGCGCCCATGATGCTTAAAGAGAGAGAGAGAGAGAGAGAGAG 2462
QY 541 CGGCAAGGAGCAAGGCTACCGTCAATGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 600
DB 2463 AGGCAAGGAGCAAGGCTACCGTCAATGATGCTTCTTCTTCTTCTTCTTCTTCTTCT 2522
QY 601 AGTAAAGGAGCAAGGCTACCGTCAATGATGCTTCTTCTTCTTCTTCTTCTTCTTCTT 660
DB 2523 AACAAAGGAGCAAGGCTACCGTCAATGATGCTTCTTCTTCTTCTTCTTCTTCTTCTT 2582
QY 661 ACAAGAGAGCAAGGCTACCGTCAATGATGCTTCTTCTTCTTCTTCTTCTTCTTCTT 720
DB 2583 ACCAAGAGAGCAAGGCTACCGTCAATGATGCTTCTTCTTCTTCTTCTTCTTCTTCT 2642
QY 721 GAGATCAAACTCACAAGAGCAAGGCTACCGTCAATGATGCTTCTTCTTCTTCTTCTTCT 780
DB 2643 GAGATCAAACTCACAAGAGCAAGGCTACCGTCAATGATGCTTCTTCTTCTTCTTCTT 2702
QY 781 AGTATGATGAG 840
DB 2703 AACGAG 2762
QY 841 ACTTCTGATGAG 900
DB 2763 ACCAG 2822
QY 901 CAACCAAG 960
DB 2823 CAGCCAG 2882
QY 961 GAACCTAGACAAATAAACAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 2883 GAGCCCAACAAG 2942
QY 1021 ATCTCTACTTAACAACCTTCCAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 2943 ATGAGACCAACAACAG 3002
QY 1081 ACCACCAATGACAAACACAG 1140
DB 3003 ACCACCAATGACAAACAG 3062
QY 1141 ATAAACAACCTGCTCCAAACGGGAAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB 3063 AAGACCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3122
QY 1201 GGGCCGGCCACAGCGGCAACAGGCAAACTGATGATGATGATGATGATGATGATGATGAT 1260
DB 3123 GGGCCGGCCACAGCGGCAACAGGCAAACTGATGATGATGATGATGATGATGATGATGAT 3182
QY 1261 CCCAGCTGCACTGCAACAACTCTGTATATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAT 1320
DB 3183 CCCAGCTGCACTGCAACAACTCTGTATATTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAT 3242
QY 1321 GAAAGTATATATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1380

```
Db 3243 GAGGGCGATATGTTCCCTTCTTGACGGCTGATCAAGCGCCCTATCGACTTCGACCCC 3302
Qy 1381 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCAGCTTTTAATACTTTCAACTAATGAG 1440
Db 3303 GTGCCCAACACAGACCATCTTCGACGAGAGCAGCAGCGGGCCGAGCGCGAGGAG 3362
Qy 1441 GAACAAACACTCCCGCGAATATCAGTTTAATCTTCTTCTTATTTCTTGATATAAAATGGA 1500
Db 3363 GACCAAGCAGCGCGAGCCCAACATCAGCTGACCTGACCTGAGCTACTTCCCAACATCAACAG 3422
Qy 1501 GATAGCTGCTACTCTGCGGGAACACAGAAATGATGATGATGATGATGATGATGATGATGATG 1560
Db 3423 AACACCGCTACAGCGGCGAGAACAGAAACGACTGCGACGCGAGCTGAGGATCTGGAGC 3482
Qy 1561 GTGCAGGAGGACGATTTGCGCGCAGGGCTTAGCTGGATACCAATTTTGGCCCTGGAATC 1620
Db 3483 GTGCAGGAGGACGACTTGGCGCGCGCTGAGCTGATTCCTTCTTCCGCCCGGCATC 3542
Qy 1621 GAAGGACTCTATACGCGCGGTTTAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG 1680
Db 3543 GAGGCGCTGTACACCGCGCGCTGATCAAGAACCAAGAACCAACCTGGTGTGCGAGGCTGAGG 3602
Qy 1681 CGCTTAGCTTAATCAAACTGCTAAATCTTGGAGCTCTTTGTTAAAGGCTCACAACCGAGAA 1740
Db 3603 AGGCTGGCCACACGACCGCGCAAGAGCTTGGAGCTGCTGAGGCTGACCAACCGAGGAG 3662
Qy 1741 AGGACATTTTCTTAATCAATAGGATGCAATTTGACTTTTGTCTTACGAGGTGGGGCGGA 1800
Db 3663 AGGACCTTCAGCCTGATCAACAGGACCGCATGAGCTTCTGCTGACCGAGGTGGGGCGG 3722
Qy 1801 ACATGCAAGGTGCTAGGACCTGATTTGTTGATAGGAATAGAGATCTATCTAAAAATATC 1860
Db 3723 ACCTGCAAGGTGCTGGGCGCGGCTGCTGCTATCGGCATCGAGGACCTGAGCAGGAAATC 3782
Qy 1861 TCAGAACAAATCGACAAATCAGAAAGATGAAACAAAGGAGGAACTGGCTGGGGTCTA 1920
Db 3783 AGCGAGCAGATCGACAGATCAAGAAAGGACGAGCAGAGGAGGCGCACCGGCTGGGGCTG 3842
Qy 1921 GGTGCAAAATGGTGACATCTGACTG 1946
Db 3843 GCGGCAAGTGGTGACCGGACTG 3868
```

RESULT 13

```
US-10-066-506A-1
; Sequence 1, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 1
; LENGTH: 2252
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Ebola virus Zaire Mayinga strain Glycoprotein 2
; OTHER INFORMATION: and Marburg virus strain Musoke Glycoprotein 2
US-10-066-506A-1
```

Query Match

```
Best Local Similarity 26.7%; Score 546.6; DB 15; Length 2252;
Matches 624; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
```

```
Qy 1294 AGRAGAAGCGATCGATTTTCTGGAAGAGGTGATATATCCCGTTTCTTAGATGGGTTA 1353
Db 1492 AGAACTCGACGATCGATCTCTGGAGGAGGAGCGACATGTTCCCTTTCTGCGGGGTTA 1551
Qy 1354 ATAATACTGAAATTTGATTTTGAATCCAAATCCCAACACAGAAACAAATCTTTGATGAATCT 1413
Db 1552 ATAATGCTCCAATTTGATTTTGACCCAGTTCCAATACAAAACAAATCTTTGATGAATCC 1611
Qy 1414 CCCAGCTTTAATACTTCAACTAAATGAGGAAACAAACACTCCCGCGAATATCAGTTTAATCT 1473
Db 1612 TCTAGTTCTGCTGCTCGGCTGAGGAAAGTCAACATGCTCCCTCCCAATATATAGTTTAATCT 1671
Qy 1474 TTCTCTTATTTTCTGATTAATAATGAGAGATCTGCTACTCTGGGGAACAGAGATCAT 1533
Db 1672 TTAATCTTTTCTTAATATAATGAGAACTGCTACTCTGAGGAAATATGAGAAATGAT 1731
Qy 1534 TGTGATGACAGATTTGAGGATTTGAGGTGTGAGGAGGACGATTTTGGCGGCGAGGCTTAGC 1593
Db 1732 TGTGATGACAGATTTAAGAAATTTGGAGCGTTTCAGGAGGATGACCTGGCGCGAGGCTCAGT 1791
Qy 1594 TGGATACCAATTTTGGCCCTTGGAAATCGAAGGACTCTATATCTGCGGGTTTAAATCAAAAAT 1653
Db 1792 TGGATACCGTTTGGCCCTTGGAAATTCAGAGGACTTTACACTGCTGTTTAAATTAATAAT 1851
Qy 1654 CAGAACAAATTTAGTTTGTGAGGCTTGGAGGCTTAGCTAATCAAACTGCTAAATCTCTTGAG 1713
Db 1852 CAAAACAAATTTGGTCTGAGGCTTGGAGGCTTAGCCAAATCAAACTGCTCAAAATCTCTTGAA 1911
Qy 1714 CTCTTGTAAAGGTCACAACCGAGGAAAGGACATTTTCTTAAATCAATAGGATGCAAT 1773
Db 1912 CTCTTATGAGATCAACAACTGAGGAAAGGACATTTCTTAAATCAATAGACATGCTATT 1971
Qy 1774 GACTTTTGTCTACAGGTGGGCGGAAACATGCAAGGTGTAGGACCTGATTTGTCATA 1833
Db 1972 GACTTTCTACTCAAGATGGGAGGAAACATGCAAGGTGTAGGACCTGATTTGTCATC 2031
Qy 1834 GGAATAGAGATCTATCTAAAATATCTCAGAACAAATCGAAGGAAATCAGAAAGGATGAA 1893
Db 2032 GGGATAGAAGACTTGTCCAAAATATTTTCAGAGCAAAATGACCAAAATTAAGAGGACGAA 2091
Qy 1894 CAAAAGGAGGAAACTGGCTGGGGTCTAGGTGGCAAAATGGTGGACATCTGACTGGGGTGT 1953
Db 2092 CAAAAGGAGGAGTGTGGTGGGGTCTGGGTGGTAAATGGTGGACATCCGACTGGGGTGT 2151
Qy 1954 CTCACCAATTTGGGCATCCTGCTACTATATATCTAGCTGTTCTGATTTGCTCTGCTCTGT 2013
Db 2152 CTTACTAACTTTGGGCAATTTGCTACTATTATCCATAGCTGTTGATTTGCTCTATCTCTGT 2211
Qy 2014 ATCTGTGATCTTCTTCACTAAATACATTTGGATGA 2046
Db 2212 ATTTGTGATCTTCTTACTAAATATATCGGATGA 2244
```

RESULT 14

```
US-10-491-121-15
; Sequence 15, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
```



```
; SEQ ID NO 15
; LENGTH: 6940
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012-GP(S) delta TM
US-10-491-121-15

Query Match          7.1%; Score 145; DB 20; Length 6940;
Best Local Similarity 62.6%; Pred. No. 3.3e-31;
Matches 226; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 1530 TGATTGTGATGAGAGTTGAGGATTTGGAGTGTGAGAGGAGGATTTGGCGGCGAGGCT 1589
Db 3483 TAAATGCAATCCAACTTACACTACTGGACTGCACAAGAAACAATAATGTCTGTGGAT 3542
QY 1590 TAGCTGGATACCATTTTGGCCCTGGATCGAAGGACTCTATCTGCGGTTTAAATCAA 1649
Db 3543 TGCCTGGATCCCGTACTTTGGACCGGGTGCAGAAGGCATATACACTGAAGGCCCTTATGCA 3602
QY 1650 AAATCAGAACAAATTTAGTTTGTAGGTTGAGGCGCTTAGCTAATCAAACTGCTAAATCCTT 1709
Db 3603 CAACCAAAATGCTTAGTCTGGACTCAGACAACTTGCNAATGAACAACACTCAAGCTCT 3662
QY 1710 GGAGCTCTTTGTAAGGTCACAAACGAGGAAAGGACATTTTCCCTTAATCAATAGGATGC 1769
Db 3663 GCAGCTTTTCTTAAGGGCCACGACGAGCTGGGACATATACCATCTCAATAGGAAGGC 3722
QY 1770 AATTGACTTTTGTACGAGTGGGCGGACATCGAAGGTCAGGCTAGGACCTGATTTG 1829
Db 3723 CATAGATTTCTTTCGCGACGATGGGCGGACATGTAGGATCCTGGGACCAAGATTGTTG 3782
QY 1830 CATAGGAATAGAGATCTATCTTAAATAATCTCAGAACAAATCGACAAATCAGAAAGGA 1889
Db 3783 CATTGAGCCACATGATTGACCAAAAACATCACTGATAAATCAACCAATCATCCATGA 3842
QY 1890 T 1890
Db 3843 T 3843

RESULT 15
US-10-860-878-2
; Sequence 2, Application US/10850878
; Publication No. US20050130129A1
; GENERAL INFORMATION:
; APPLICANT: Nabel, Gary J
; APPLICANT: Delgado, Rafael
; APPLICANT: Yang, Zhi-yong
; TITLE OF INVENTION: Targeting Gene Transfer Vectors To Certain Cell Types
; FILE REFERENCE: 2115a-001474POA
; CURRENT APPLICATION NUMBER: US/10/860,878
; CURRENT FILING DATE: 2004-06-03
; PRIOR APPLICATION NUMBER: US/09/600,766
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: USSN 60/072033
; PRIOR FILING DATE: 1998-01-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 7073
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Plasmid
; OTHER INFORMATION: containing DNA for GP of Ebola Virus, Sudan strain
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (886)..(1129)
; OTHER INFORMATION: CMV IE 5' UT
; FEATURE:
; NAME/KEY: Intron
```

```
; LOCATION: (1130)..(1840)
; OTHER INFORMATION: CMV IE INT
; FEATURE:
; NAME/KEY: enhancer
; LOCATION: (248)..(885)
; OTHER INFORMATION: CMV enhancer
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1870)..(4089)
; OTHER INFORMATION: GP(S)
; FEATURE:
; NAME/KEY: gene
; LOCATION: (4090)..(4542)
; OTHER INFORMATION: TDGH
; FEATURE:
; NAME/KEY: gene
; LOCATION: (5138)..(6760)
; OTHER INFORMATION: Kan r
US-10-860-878-2
```

```
Query Match          7.1%; Score 145; DB 22; Length 7073;
Best Local Similarity 62.6%; Pred. No. 3.4e-31;
Matches 226; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 1530 TGATTGTGATGAGAGTTGAGGATTTGGAGTGTGAGAGGAGGATTTGGCGGCGAGGCT 1589
Db 3503 TAAATGCAATCCAACTTACACTACTGGACTGCACAAGAAACAATAATGTCTGTGGAT 3562
QY 1590 TAGCTGGATACCATTTTGGCCCTGGATCGAAGGACTCTATCTGCGGTTTAAATCAA 1649
Db 3563 TGCCTGGATCCCGTACTTTGGACCGGGTGCAGAAGGCATATACACTGAAGGCCCTTATGCA 3622
QY 1650 AAATCAGAACAAATTTAGTTTGTAGGTTGAGGCGCTTAGCTAATCAAACTGCTAAATCCTT 1709
Db 3623 CAACCAAAATGCTTAGTCTGGACTCAGACAACTTGCNAATGAACAACACTCAAGCTCT 3682
QY 1710 GGAGCTCTTTGTAAGGTCACAAACGAGGAAAGGACATTTTCCCTTAATCAATAGGATGC 1769
Db 3683 GCAGCTTTTCTTAAGGGCCACGACGAGCTGGGACATATACCATCTCAATAGGAAGGC 3742
QY 1770 AATTGACTTTTGTACGAGTGGGCGGAAACATGCAAGGTCGTAGGACCTGATTTG 1829
Db 3743 CATAGATTTCTTCTGCGACGATGGGCGGACATGTAGGATCCTGGGACCAAGATTGTTG 3802
QY 1830 CATAGGAATAGAGATCTATCTTAAATAATCTCAGAACAAATCGACAAATCAGAAAGGA 1889
Db 3803 CATTGAGCCACATGATTGGACCAAAAACATCACTGATAAATCAACCAATCATCCATGA 3862
QY 1890 T 1890
Db 3863 T 3863
```

Search completed: August 1, 2005, 20:41:05
Job time : 1291.84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 31, 2005, 13:57:58 ; Search time 8793.18 Seconds
(without alignments)
11274.581 Million cell updates/sec

Title: US-10-066-506A-7
Perfect score: 2046
Sequence: 1 atgaagaccatatattttct.....ttactaatatateggataa 2046

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_in.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pi.*
9: gb_pi.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2046	100.0	2046	6	AX717719	Sequence
2	1847.6	90.3	2046	6	AX717725	Sequence
3	1842.8	90.1	2046	14	AF005734	Marburg v
4	1514.8	74.0	11460	6	AF209577	Sequence
5	1514.8	74.0	19104	14	WVREPCYC	212132 Marburg vir
6	1513.2	74.0	19112	14	AY430366	AY430366 Lake Vict
7	1513.2	74.0	19113	14	AY430365	AY430365 Lake Vict
8	1485.6	72.6	2051	6	AX717721	Sequence
9	1474.8	72.1	2046	14	AF005735	Marburg v
10	1473.2	72.0	2948	14	MAVSPA	68493 Marburg Vir
11	1473.2	72.0	19112	14	WVIRPR	29337 Marburg vir
12	1443.4	70.5	19151	14	AY358025	Marburg v
13	1441.2	70.4	2046	14	AF005733	Marburg v
14	1319.6	64.5	2046	6	AX717717	Sequence
15	780	38.1	1841	6	AX717715	Sequence
16	746.6	36.5	2252	6	AX717713	Sequence
17	149	7.3	18890	14	AB050936	Reston Eb
18	148.2	7.2	2408	14	EVU28006	Cote d'Ivoi
19	145.8	7.1	2033	14	EVU23416	Reston Ebol

20	145.8	7.1	2033	14	EVU23417	U23417 Reston Ebol
21	145.8	7.1	2233	14	AF034645	Ebola vir
22	145.8	7.1	2360	14	EVU23152	U23152 Reston Ebol
23	145.8	7.1	18891	14	AF522874	AF522874 Reston Eb
24	139.2	6.8	2261	14	AY344234	AY344234 Sudan ebo
25	139.2	6.8	2363	14	AY316199	AY316199 Sudan Ebo
26	139.2	6.8	18875	14	AY729654	AY729654 Sudan ebo
27	137.2	6.7	2408	14	EVU81161	U81161 Zaire Ebola
28	135.6	6.6	18959	14	AY142960	AY142960 Zaire Ebo
29	135.6	6.6	18960	14	AF499101	AF499101 Zaire Ebo
30	134.4	6.6	2362	14	EVU23069	U23069 Sudan Ebola
31	134.4	6.6	2362	14	EVU28134	U28134 Sudan Ebola
32	134	6.5	2039	6	AX717723	AX717723 Sequence
33	134	6.5	2298	6	AR404895	AR404895 Sequence
34	134	6.5	2298	6	AX092102	AX092102 Sequence
35	134	6.5	2406	14	EVU31033	U31033 Zaire Ebola
36	134	6.5	2408	14	EVU23187	U23187 Zaire Ebola
37	134	6.5	11742	14	EBORNA	L11365 Zaire Ebola
38	134	6.5	18959	14	AF086833	AF086833 Zaire Ebo
39	134	6.5	18959	14	AF272001	AF272001 Zaire Ebo
40	130.8	6.4	2408	14	EVU28077	U28077 Zaire Ebola
41	130.8	6.4	18961	14	AY354458	AY354458 Zaire ebo
42	127.6	6.2	2173	14	AY058898	AY058898 Zaire Ebo
43	126	6.2	2173	14	AY526098	AY526098 Zaire ebo
44	126	6.2	2173	14	AY526100	AY526100 Zaire ebo
45	126	6.2	2173	14	AY526101	AY526101 Zaire ebo

ALIGNMENTS

RESULT 1
AX717719
LOCUS AX717719 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 7 from Patent WO02079239.
ACCESSION AX717719
VERSION AX717719.1 GI:29890728
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.

REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 7 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)

FEATURES
source
1..2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chimeric molecule between Marburg virus strain
Raven Glycoprotein 1 and Marburg virus strain Musoke
Glycoprotein 2"

ORIGIN

Query Match 100.0%; Score 2046; DB 6; Length 2046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTATCCAAAGTATATAAACTCTCCCT 60
Db 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTATCCAAAGTATATAAACTCTCCCT 60
Qy 61 GTTTTAGAAATTCGTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTCTCCGGAACC 120
Db 61 GTTTTAGAAATTCGTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTCTCCGGAACC 120
Qy 121 CTCCAAAGACAGAAAGATGTTTCATCTCATGATGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCCAAAGACAGAAAGATGTTTCATCTCATGATGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Qy 181 GATTCCCTTTTGGGAAGCATCTAAAGATGGGCTTTCAGGACAGGTGTTCTCTCCCAAGAAC 240

Db 181 GATTCCTCCCTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGGTTCTCTCCCAAGAAC 240
QY 241 GTTGGATATACGGAAGGAGAGCAAGCCAAACATGTTACAATATAAAGTGAACAGACCT 300
Db 241 GTTGGATATACGGAAGGAGAGCAAGCCAAACATGTTACAATATAAAGTGAACAGACCT 300
QY 301 TCTGAAATCTTGCTGCTGGATCCTCCAGTAAATATCCGCGATATACCTAAATGTAAA 360
Db 301 TCTGAAATCTTGCTGCTGGATCCTCCAGTAAATATCCGCGATATACCTAAATGTAAA 360
QY 361 ACTGTTCAATATATCAAGGTCAAAACCTCATGACAGGGGATGCGCTCATTTGTGG 420
Db 361 ACTGTTCAATATATCAAGGTCAAAACCTCATGACAGGGGATGCGCTCATTTGTGG 420
QY 421 GGGGCAATTTCTTGATGATCGCTGCTCTCAACAATGTACCGAGGCAAGGTCTTC 480
Db 421 GGGGCAATTTCTTGATGATCGCTGCTCTCAACAATGTACCGAGGCAAGGTCTTC 480
QY 481 ACTGAAGGAATATATAGCAGCTATGATTTGTTAATAAGACAGTTTCACAGAATGATTTTTCT 540
Db 481 ACTGAAGGAATATATAGCAGCTATGATTTGTTAATAAGACAGTTTCACAGAATGATTTTTCT 540
QY 541 AGGCAAGGACAAAGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTTGGACAAGC 600
Db 541 AGGCAAGGACAAAGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTTGGACAAGC 600
QY 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTTGGCATCTCCCAAGAATACAACTCC 660
Db 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTTGGCATCTCCCAAGAATACAACTCC 660
QY 661 ACAACAATCAAAATGATGATGCTCTTAAACCTCCATCCCTCCCAACAGTAAGTCCG 720
Db 661 ACAACAATCAAAATGATGATGCTCTTAAACCTCCATCTTAAACCTCCATCCCTCCCAACAGTAAGTCCG 720
QY 721 AGCATTCATCTPACAAATATCAAAATTAATCTGCTAAATCTGGAACCTATGAACCCAAGT 780
Db 721 AGCATTCATCTPACAAATATCAAAATTAATCTGCTAAATCTGGAACCTATGAACCCAAGT 780
QY 781 AGCGAGATGAGGACCTTATGATTTCCGCTCAGGATCTGGAGACAGGGGCCACACA 840
Db 781 AGCGAGATGAGGACCTTATGATTTCCGCTCAGGATCTGGAGACAGGGGCCACACA 840
QY 841 ACTCTTAATGTAGTCACTGAAACAGAAACAACTCGTCAACAATATTTGTCACCTCTCACTA 900
Db 841 ACTCTTAATGTAGTCACTGAAACAGAAACAACTCGTCAACAATATTTGTCACCTCTCACTA 900
QY 901 CATCAAGACCTCAACATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAAC 960
Db 901 CATCAAGACCTCAACATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAAC 960
QY 961 GAGCAATGGAACCGACCAACACACACAGCAACGCTCTCAACAATACTAATACA 1020
Db 961 GAGCAATGGAACCGACCAACACACACAGCAACGCTCTCAACAATACTAATACA 1020
QY 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCGCAAC 1080
Db 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCGCAAC 1080
QY 1081 ATCAACAATATATGATACAAAGTGAATAGCAGAAAGGCAACAAACCAATGCTCAGTTG 1140
Db 1081 ATCAACAATATATGATACAAAGTGAATAGCAGAAAGGCAACAAACCAATGCTCAGTTG 1140
QY 1141 AACACAATCTAGATCCCAAGAAATCCCAACAGCAGACAGACCAACAGCAGCAAC 1200
Db 1141 AACACAATCTAGATCCCAAGAAATCCCAACAGCAGACAGACCAACAGCAGCAAC 1200
QY 1201 AACATCATCATGACATCATGATATACAGCAACACCCCAACAAATTTCTCCGAT 1260
Db 1201 AACATCATCATGACATCATGATATACAGCAACACCCCAACAAATTTCTCCGAT 1260
QY 1261 TCTAGTCCGACAAACCGGCTCTCTATATATCTTTAGAAAGAACGATCGATCTCTGGAGG 1320

Db 1261 TCTAGTCCGACAAACCGGCTCTCTATATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
QY 1321 GAAGCGCACATGTTCCCTTTCTCGATGGGTAAATAATGCTCCAAATGATTTTGGACCA 1380
Db 1321 GAAGCGCACATGTTCCCTTTCTCGATGGGTAAATAATGCTCCAAATGATTTTGGACCA 1380
QY 1381 GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTTCGTGCTCGGCTGAGGAA 1440
Db 1381 GTTCCAAATACAAAAACAATCTTTGATGAATCTCTAGTTCGTGCTCGGCTGAGGAA 1440
QY 1441 GATCAACATGCTCCGCCAATATTAGTTTAACCTTTATCTTATTTTCTTAATAATAATGAG 1500
Db 1441 GATCAACATGCTCCGCCAATATTAGTTTAACCTTTATCTTATTTTCTTAATAATAATGAG 1500
QY 1501 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGATGAGAGAGTTAAGAAATTTGGAGC 1560
Db 1501 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGATGAGAGTTAAGAAATTTGGAGC 1560
QY 1561 GTTCAGAGGATGACCTGGCGCAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAT 1620
Db 1561 GTTCAGAGGATGACCTGGCGCAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAT 1620
QY 1621 GAAGGACTTTACCTGCTGTTTTTAATAATAATAATAATAATAATAATAATAATAATA 1680
Db 1621 GAAGGACTTTACCTGCTGTTTTTAATAATAATAATAATAATAATAATAATAATAATA 1680
QY 1681 CGTCTAGCCAAATCAAACTGCCAAATCTTTGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
Db 1681 CGTCTAGCCAAATCAAACTGCCAAATCTTTGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
QY 1741 AGAACATTTCTCTTAATCAATAGACATGCTATTGATGCTTACTCAACAAGATGGGAGGA 1800
Db 1741 AGAACATTTCTCTTAATCAATAGACATGCTATTGATGCTTACTCAACAAGATGGGAGGA 1800
QY 1801 ACATGCAAAAGTCTTGACCTGATTTGCAATCGGATAGAGAGTCTGTCACAAATATTT 1860
Db 1801 ACATGCAAAAGTCTTGACCTGATTTGCAATCGGATAGAGAGTCTGTCACAAATATTT 1860
QY 1861 TCAGAGCAAAATTTGACCAAAATTTAAAGAGGAGGAGCTGGTTGGGCTCTG 1920
Db 1861 TCAGAGCAAAATTTGACCAAAATTTAAAGAGGAGGAGCTGGTTGGGCTCTG 1920
QY 1921 GGTGTAATGTTGGACATCCGACTGGGCTTTCTTAACTTTGGGCAATTTTGGCTACTA 1980
Db 1921 GGTGTAATGTTGGACATCCGACTGGGCTTTCTTAACTTTGGGCAATTTTGGCTACTA 1980
QY 1981 TTATCCATAGCTGCTTGAATGCTCTATCTGATTTTGTCTATTTTGTCTATTTTACTAATATATC 2040
Db 1981 TTATCCATAGCTGCTTGAATGCTCTATCTGATTTTGTCTATTTTGTCTATTTTACTAATATATC 2040
QY 2041 GGATAA 2046
Db 2041 GGATAA 2046

RESULT 2
AX717725
LOCUS AX717725 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 13 from Patent WO02079239.
ACCESSION AX717725
VERSION AX717725.1 GI:29890731
KEYWORDS Lake Victoria marburgvirus
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
Virus; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1
AUTHORS Grogan, C.C., Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 13 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source Location/Qualifiers
1 .2046

```

/organism="Lake Victoria marburgvirus"
/mol_type="unassigned DNA"
/strains="Raven"
/db_xref="caxon:11269"
/!note="chimeric molecule between Marburg virus
Glycoprotein 1 and Marburg virus Glycoprotein 2"

ORIGIN
Query Match      90.3%; Score 1847.6; DB 6; Length 2046;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 1922; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTCTGATTAGTCTCTATTTTAAATCCAAAGTATATAAACTCTCCCT 60
DB 1 ATGAAGACCATATATTTCTGATTAGTCTCTATTTTAAATCCAAAGTATATAAACTCTCCCT 60
QY 61 GTTTTAAAGAAATGCTAGTAACAGCCAACTCAAGATGTAGATTTCAGTGTGCTCCGGAAAC 120
DB 61 GTTTTAAAGAAATGCTAGTAACAGCCAACTCAAGATGTAGATTTCAGTGTGCTCCGGAAAC 120
QY 121 CTCMAAGACAGAAGATGTTTCTATCTGATGGATTTTACATGATGGGCAAAAAGTTGCT 180
DB 121 CTCMAAGACAGAAGATGTTTCTATCTGATGGATTTTACATGATGGGCAAAAAGTTGCT 180
QY 121 CTCMAAGACAGAAGATGTTTCTATCTGATGGATTTTACATGATGGGCAAAAAGTTGCT 180
DB 121 CTCMAAGACAGAAGATGTTTCTATCTGATGGATTTTACATGATGGGCAAAAAGTTGCT 180
QY 181 GATTTCCCTTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGATGTTTCTCCCAAGAAC 240
DB 181 GATTTCCCTTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGATGTTTCTCCCAAGAAC 240
QY 241 GTTGAAGTACGGAAGGAGAGAGCAAAACATGTTTACAATATAAGTGTAAACAGACCT 300
DB 241 GTTGAAGTACGGAAGGAGAGAGCAAAACATGTTTACAATATAAGTGTAAACAGACCT 300
QY 301 TCTGMAAATCTTCTGCTGCTGATCTCCAGTAAATTCGCGGATATACCTTAATATGTA 360
DB 301 TCTGMAAATCTTCTGCTGCTGATCTCCAGTAAATTCGCGGATATACCTTAATATGTA 360
QY 361 ACTGTTTCATATATTCAGGTCATAAAGCTCTATGACAGAGGATGTTCCCTCCATTTG 420
DB 361 ACTGTTTCATATATTCAGGTCATAAAGCTCTATGACAGAGGATGTTCCCTCCATTTG 420
QY 421 GGGGCAATTTTCTGTATGATCGGCTGCTCTCAACAATATGACAGAGGATGTTCCCTCCATTTG 480
DB 421 GGGGCAATTTTCTGTATGATCGGCTGCTCTCAACAATATGACAGAGGATGTTCCCTCCATTTG 480
QY 481 ACTGAAGGAAATATAGCAGCTATGATGTTTAAAGACAGATTCACAGAATGATTTTCT 540
DB 481 ACTGAAGGAAATATAGCAGCTATGATGTTTAAAGACAGATTCACAGAATGATTTTCT 540
QY 541 AGGCAAGCAGAGGTTATCGTCACATGAACCTTGACCTCCACCAATATATGACAGAC 600
DB 541 AGGCAAGCAGAGGTTATCGTCACATGAACCTTGACCTCCACCAATATATGACAGAC 600
QY 601 AGCAATGAAACGACAGAGAAATGATACGGGATGTTTGGCATCTCTCAAGAATACAACTCC 660
DB 601 AGCAATGAAACGACAGAGAAATGATACGGGATGTTTGGCATCTCTCAAGAATACAACTCC 660
QY 661 ACAACAAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACCTCG 720
DB 661 ACAACAAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACCTCG 720
QY 721 AGCATTTCACTTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
DB 721 AGCATTTCACTTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 780
QY 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCACACACA 840
DB 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCACACACA 840
QY 841 ACTCTTAATGATGATCACTGAACAGAGAAACAACTCGTCAACAATATGTTCCACTCTTCACTA 900
DB 841 ACTCTTAATGATGATCACTGAACAGAGAAACAACTCGTCAACAATATGTTCCACTCTTCACTA 900
QY 901 CATCCAGACCTCAACATGACGAGCAAAACAGTACGAATCTCTCCCGACATGCTGTAAC 960
```

```

DB 901 CATCCAGCACCTCACAACATGAGCAAAACAGTACGATCTCTCCGACATGCTGTAAC 960
QY 961 GAGCACAATGGAACCGACCCCAACAAACAACAGCAACAGCTCTCTCAACAATACATAACA 1020
DB 961 GAGCACAATGGAACCGACCCCAACAAACAACAGCAACAGCTCTCTCAACAATACATAACA 1020
QY 1021 ACTCCGACCTTAACACTCTCAAGTCAACCTCAAGTACTCTTCCCTCCGACCGGAAC 1080
DB 1021 ACTCCGACCTTAACACTCTCAAGTCAACCTCAAGTACTCTTCCCTCCGACCGGAAC 1080
QY 1081 ATCAACCAATATGATATACAACAGTGAATAGCAGAAAGCAACAAACCAATGCTCAGTTG 1140
DB 1081 ATCAACCAATATGATATACAACAGTGAATAGCAGAAAGCAACAAACCAATGCTCAGTTG 1140
QY 1141 AACCAACTCTTAGATCCAAACAGAAATATCCCAACAGGACAGACACCAACAGCAACACC 1200
DB 1141 AACCAACTCTTAGATCCAAACAGAAATATCCCAACAGGACAGACACCAACAGCAACACC 1200
QY 1201 AACATCATCATGACGACATCAGATATTAACAGCAAAACCCCAACCAATTTCTTCCGGAT 1260
DB 1201 AACATCATCATGACGACATCAGATATTAACAGCAAAACCCCAACCAATTTCTTCCGGAT 1260
QY 1261 TCTAGTCCGACAAACCCCTCTCTATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
DB 1261 TCTAGTCCGACAAACCCCTCTCTATATCTTTAGAAAGAAACGATCGATTTCTGGAAA 1320
QY 1321 GAAGGCGACATGTTCCCTTTTCTGGATGGGTTAAATAATGCTCCAATTTGATTTTGACCCA 1380
DB 1321 GAAGGCGATATATTTCCCGTTTTAGATGGGTTAAATAATCTGAAAATTTGATTTGATCCA 1380
QY 1381 GTTCCAAATACAAAACAACTTTTATGATGAATCTCTAGTCTGGTGGCTCCGCTGAGGAA 1440
DB 1381 ATCCCAACACAGAAACAACTTTTATGATGAATCTCTAGTCTGGTGGCTCCGCTGAGGAA 1440
QY 1441 GATCAACATGCTCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTAAATAAATGAG 1500
DB 1441 GAAACAACATCTCCCGAATATCAGTTTAACTTTTCTTATTTTCTGATAAANAATGGA 1500
QY 1501 AACACTGCTACTCTCGGAGAAATGAGAAATGATTTGATGACAGAGTTAAGAAATTTGGAGC 1560
DB 1501 GATACTGCTACTCTCGGAGAAACGAGAAATGATTTGATGACAGAGTTGAGGATTTGGAGT 1560
QY 1561 GTTCAGGAGGATGACCTGGCCGACAGGCTCAGTTGGATACCGTTTTTTGGCCCTCGGAAT 1620
DB 1561 GTGAGGAGGACGATTTGGCCGACAGGCTTAGCTGGATACCAATTTTTGGCCCTCGGAAT 1620
QY 1621 GAAGGACTTTACACTGCTGTTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1680
DB 1621 GAAGGACTTTATATCTGCGGTTTTAATAAATAAATAAATAAATAAATAAATAAATAAATA 1680
QY 1681 CGTCTAGCAATCAAACTGCCAAATCTCTTGAACTCTTTATTTGAGAGTCACAACTGAGGAA 1740
DB 1681 CGCTTAGCTAATCAAACTGCTAATCTTTGAGCTCTTTGTAAGGTCACAACTGAGGAA 1740
QY 1741 AGAACAATCTCTTAATCAATAGACATGCTATGATGCTTTCTACTCAAGATGGGGAGA 1800
DB 1741 AGGACATTTTCTTAAATCAATAGGACGCAATGACATTTTGTCTTACGAGGTGGGGCGGA 1800
QY 1801 ACATGCAAGTCTTGGACCTGATTTGTCATCGGATAGAGACTCTGTCGCAAAATAAT 1860
DB 1801 ACATGCAAGTCTTGGACCTGATTTGTCATCGGATAGAGACTCTGTCGCAAAATAAT 1860
QY 1861 TCAGAGCAAAATGACCAAAATTAATAAAGGACCAACAAAGAGGAGGACTGGTTGGGGCTG 1920
DB 1861 TCAGACAAATCGACAAATTCAGAAAGGATGACAAAGAGGAGGAACTGGCTGGGGCTA 1920
QY 1921 GGTGGTAAATGGTGGACATCCGACTGGGGTCTTCTACTAATCTGGGCAATTTGCTACTA 1980
DB 1921 GGTGGCAAAATGGTGGACATCTGACTGGGGTCTTCTCAACCAATTTGGGCACTCTGCTACTA 1980
QY 1981 TTATCCATAGCTGCTTGAATGCTCTATCTCTGATTTGTCGTAATTTGCTTACTTAAATATC 2040
```

Db	1981	TTATCTATAGCTGTTCTGATTGCTCTGCTGCTATCTATCTGCTGATCTTCACTAAATACATT	2040
QY	2041	GGATAA 2046	
Db	2041	GGATGA 2046	
RESULT 3			
AF005734			
LOCUS			
DEFINITION	AF005734	2046 bp RNA linear VRL 23-JAN-1998	
ACCESSION	AF005734	Marburg virus strain M/Kenya/Kitum Cave/1987/Ravn glycoprotein precursor (GP) gene, complete cds.	
VERSION	AF005734.1	GI:2459877	
KEYWORDS			
SOURCE			
ORGANISM		Lake Victoria marburgvirus	
REFERENCE		Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.	
AUTHORS		Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and Feldmann,H.	
TITLE		Variation in the glycoprotein and VP35 genes of Marburg virus strains	
JOURNAL		Virology 240 (1), 138-146 (1998)	
MEDLINE		98110148	
PUBMED		9448698	
REFERENCE		2 (bases 1 to 2046)	
AUTHORS		Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.	
TITLE		Direct Submission	
JOURNAL		Submitted (29-MAY-1997) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Building 15, Room SB611, Mail Stop G14, Atlanta, GA 30333, USA	
FEATURES		Location/Qualifiers	
source		1..2046	
		/organism="Lake Victoria marburgvirus"	
		/mol_type="genomic RNA"	
		/strain="M/Kenya/Kitum Cave/1987/Ravn"	
gene		/db_xref="taxon:11269"	
		1..2046	
		/gene="GP"	
CDS		1..2046	
		/gene="GP"	
		/note="virion surface glycoprotein"	
		/codon_start=1	
		/product="glycoprotein precursor"	
		/protein_id="AAC40459.1"	
		/db_xref="GI:2459878"	
		/translation="MKTIYFLISLILIOSIKTLPVLEIASNSQPDVDSVCSGLQKT EDVHLMGFTLSGQKVDSPLEASKRWAFRTGVPVKNVEYTEGEAKTCYNISVDPDG KSLIDPPSNIRDYPKCTVHHIOGNPHAQGIALHMGAFFLYDRVASTMYRGKVF TEGNTAAMI VNKTVHRMIFSRQGGYRHMNLTSNKYWTSSNETORDNTGCFGILQEV NSTNNQTCPSLKPSPLETPVPSIHSNTQINTAKSGTNPSPSDDEDLMSGSGEQ GPHTLVNVVQKQSSTILSPSLPSPSQHEQNSTPNSRHVTEHNGTDFTTPAIL LNNNTTPTNYLNKLTSPPTTNNNDTQRELAESQNTAQLNLTLDPTENPT GQDNTSTNIIMTSDITSKPTNSSPDSTTPPIYFRKRSIFWKEGDIFPFLDG LINTIDPPIPNTEITIDESPSFNTSNEQHTPPNISLTFSPDPKNGDTPVSGEN ENDCDAELRINSVQEDDLAAGLSWIPFPGPIEGLYTAGLKNQNNLVCRLLRANOT AKSLELLRVTEERTFSLNRHAIIDFLLRWGTCVKVLPDCCIGIEDLSKNISEQI DKIRDEQKEETGIGGKRWTSMDGVLTNLGILLLSIAVLIALUSLCIRFTKIYG"	
		1..81	
sig_peptide		/gene="GP"	
mat_peptide		82..2043	
		/gene="GP"	
misc_feature		/product="glycoprotein"	
		1930..2019	
		/gene="GP"	
		/note="encodes transmembrane anchor peptide"	
ORIGIN			
Query Match		90.1%; Score 1842.8; DB 14; Length 2046;	
Best Local Similarity		93.8%; Pred. No. 0;	
Matches 1919; Conservative		0; Mismatches 127; Indels 0; Gaps 0;	

QY 1081 ATCCAAATATGATACACACGCTGAACCTAGCAGAAAGCGAACCAAAACCAATGCTCAGTTG 1140
Db 1081 ATCCAAATATGATACACACGCTGAACCTAGCAGAAAGCGAACCAAAACCAATGCTCAGTTG 1140
QY 1141 AACCAACTCTAGATCCAAACAGAAATCCCAACAGCAAGACACCAACAGCACAAC 1200
Db 1141 AACCAACTCTAGATCCAAACAGAAATCCCAACAGCAAGACACCAACAGCACAAC 1200
QY 1201 AACATCATGACGACATCAGATATAACAAGCAAAACACCCACCAAAATCTTCTCCGAT 1260
Db 1201 AACATCATGACGACATCAGATATAACAAGCAAAACACCCACCAAAATCTTCTCCGAT 1260
QY 1261 TCTAGTCGACAAACCCGCTCTCTATATACCTTTAGAAAGAAACGATCGATCTCTGGAG 1320
Db 1261 TCTAGTCGACAAACCCGCTCTCTATATACCTTTAGAAAGAAACGATCGATCTCTGGAG 1320
QY 1321 GAAGCGACATGTTCCCTTTTCTGGATGGGTTAATAATGCTCCAATGATTTTGACCCA 1380
Db 1321 GAAGGTGATATATCCCGTTTATAGATGGGTTAATAATACTGAAATTTGATTTGATCCA 1380
QY 1381 GTTCCAAATACAAAACAACTTTGATGAATCCTCTAGTTCTGGTGCTCGGCTGAGGAA 1440
Db 1381 ATCCAAACACAGAAACAACTTTGATGAATCCTCCAGCTTTAATACTTCAACTAATGAG 1440
QY 1441 GATCAACATGCTCCCCCAATATAGTTTAACTTTATCTTATTTTCTTAATATAATGAG 1500
Db 1441 GAACAACACATCCCCCGGAATATCAGTTTAACTTTCTCTTATTTCTTGATATAAATGGA 1500
QY 1501 AACCTGCTACTCTGGAGAAATAGAAATGATGTGATGACAGAGTTAAGAAATTTGGAGC 1560
Db 1501 GATCTGCTACTCTGGGGAACACGAGATGATGTGATGACAGAGTTGAGATTTGGAGT 1560
QY 1561 GTTCAGGAGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 1620
Db 1561 GTTCAGGAGGACGATTTGGCGCAGGGCTTAGCTGGATACCAATTTTGGCCCTGGAATC 1620
QY 1621 GAAGGACTTTACATGCTGTTTTTAAATAAATCAAAACAAATTTGCTGAGGTTGAGG 1680
Db 1621 GAAGGACTTTACATGCTGTTTTTAAATAAATCAAAACAAATTTGCTGAGGTTGAGG 1680
QY 1681 CGCTAGCAATCAAACTGCCAAATCCCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
Db 1681 CGCTAGCTAATCAAACTGCTAATCTTGGAGCTCTTGTAAAGGTACAAACCGAGGAA 1740
QY 1741 AGAACATTTCTTAATCAATAGACATGCTATGCTTTCTAATCAAGATGGGAGGA 1800
Db 1741 AGGACATTTCTTAATCAATAGGATGCAATGACTTTTGTCTTACGAGGTGGGGCGA 1800
QY 1801 ACATGCAAGTCTTGGACCTGATTTGTCATCGGATAGAGACTTGTCCAAAATATT 1860
Db 1801 ACATGCAAGTCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAAAATATC 1860
QY 1861 TCAGAGCAATTTGACCAATTAAGAAAGCAACAAAGAGGGGACTGTTGGGGCTG 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGATGAACAAAGAGGAACTGGCTGGGGCTA 1920
QY 1921 GGTGTAATGTTGGACATCCGACTGGGGTGTCTTACTAATCTGGGCAATTTGCTACTA 1980
Db 1921 GGTGCAATGTTGGACATCTGACTGGGGTGTCTTCAACCAATTTGGGCACTCTGCTACTA 1980
QY 1981 TTATCCATAGCTGCTTGAATGCTCTATCCGTGATTTGCTGATTTCTTACTAATAATATC 2040
Db 1981 TTATCTATAGCTGTTCTGATTTGCTCTGCTGATCTCTGCTGATCTTCACTAAATATC 2040
QY 2041 GGATGA 2046
Db 2041 GGATGA 2046

RESULT 4

AR279677

LOCUS

DEFINITION Sequence 1 from patent US 6517842.

11460 bp

DNA

linear

PAT 10-APR-2003

ACCESSION AR279677
VERSION AR279677.1 GI:29714600
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 11460)
AUTHORS Hevey,M.C., Negley,D.L., Pushko,P., Smith,J.F. and Schmaljohn,A.L.
TITLE Marburg virus vaccines
JOURNAL Patent: US 6517842-A 1 11-FEB-2003;
FEATURES Location/Qualifiers
1..11460
source /organism="unknown"
ORIGIN /mol_type="genomic DNA"

Query Match 74.0%; Score 1514.8; DB 6; Length 11460;

Best Local Similarity 83.8%; Pred. No. 0;

Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 1 ATGAAGACCAATATATTTTCTGATTTAGTCTCAATTTTAATCCAAAGTATATAAAACTCTCCCT 60
Db 5940 ATGAAGACCAATATATTTTCTGATTTAGTCTCAATTTTAATCCAAAGTATATAAAACTCTCCCT 5999
QY 61 GTTTTAGAAATTTCTAGTAAACAGCCCACTCAAGATGTAGATTTCAATGTTGCTCCGGAAC 120
Db 6000 ATTTTGAAGATAGTCTAGTAAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAC 6059
QY 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 6060 CTCCAAAGACAGAAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 6119
QY 181 GATTCCCTTTTGGAAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 6120 GATTCCCTTTTGGAGGATCTCAAGCGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 6179
QY 241 GTTCAGTATACGGAAGGAGAGAGCCCAACATGTTACATATAAGTGTACAGACCT 300
Db 6180 GTTCAGTATACGGAAGGAGAGAGCCCAACATGTTACATATAAGTGTACAGACCT 6239
QY 301 TCTGCAAAATCCTTGTGCTGGATCTCCCAAGTAAATCCCGATTTACCTTAAATGTAAA 360
Db 6240 TCTGCAAAATCCTTGTGCTGGATCTCCCAAGTAAATCCCGATTTACCTTAAATGTAAA 6299
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGCAAGGGGTTGGCTCCATTTGFG 420
Db 6300 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCAAGGGGTCGCCCTTCATTATAG 6359
QY 421 GGGGCATTTTCTTGATGATCGCGTTGCTCTACAACAATGTACCGAGGCAAGTCTTC 480
Db 6360 GGGGCATTTTCTTGATGATCGCGTTGCTCTACAACAATGTACCGAGGCAAGTCTTC 6419
QY 481 ACTCAAGCAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCAGAGATGATTTTTTCT 540
Db 6420 ACTCAAGCAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCAGAGATGATTTTTTCT 6479
QY 541 AGGCAAGGACAAAGTTATCGTCACATGAACCTTGACCTCCCAACATATAATTGGACAAGC 600
Db 6480 AGGCAAGGACAAAGTTATCGTCACATGAACCTTGACCTCCCAACATATAATTGGACAAGT 6539
QY 601 AGCAATGAACGACAGAGAAATGATACGGGATGTTTTGGCATCTCCCAAGATATCAACTCC 660
Db 6540 AGTAAACGAAACGCAAAACGAATGACATGATGTTTTTCGGCGCTCTTCAAGAAATACAA 6599
QY 661 ACAAACATCAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAACTCCG 720
Db 6600 ACAAACATCAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAACTCCG 6659
QY 721 AGCAATCACTCTACAAATATCTCAAAATTAATACTGCTAAATCTGGAACTATGAACCCAA 780
Db 6660 GAGATCAAACTCAACAGCACCCTCAACTGATGCGCAACCAAACTCAATACCAAGCCCAAG 6719
QY 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGGCCCCACACA 840

6720	Db	 AGTGATGATGAGGACCTCGCAACATCGGGCTCAGGGTCGGGAGAACGAGAACCCACACA	6779
841	Qy	ACTCTTAATGTAGTCACTGAAACAGAAACAATCGTCAACAATATTTGTTCCACTCTCTTCACTA	900
6780	Db	ACTTCTGATGGGTTCACCAAGCAAGGCTTTTCATCAACAATGCGCACCCACTCTCTCACCA	6839
901	Qy	CATCTCAAGCACCTCAACAATGAGCAAAAAGTACGAATCTCTTCCGACATGCTGTAACT	960
6840	Db	CAACCAAGCAGCCACAGCAAGGAGGAAAACAACAAACCATTTCCTCCAAAGATGCTGTGACT	6899
961	Qy	GAGCACAAATGGAAACCGACCCCAACAACAACCAAGCAAGCTCTCTCCCTCCCAACCCGCAAC	1020
6900	Db	GAATAGACAAAAATAACACAACCTGCAACACCGTCCATGCGCCCTCATTAACACTACCA	6959
1021	Qy	ACTCCCACTTATAACACTCTCAAGTACAACTCAGTACTCTCTTCCCTCCCAACCCGCAAC	1080
6960	Db	ATCTCTACTTAACAACACTCTCAAAACAACATTTAGCACTCTCTCTGACACATTTACAAAAC	7019
1081	Qy	ATCACCAATTAATGATACAAACCGTGAACTAGCAGAAAGCGAAACAAACCAATGCTCAGTTG	1140
7020	Db	ACCACCAATGACAACACACAGAGCACAACTCACTGAAAATGAGCAAAACGAGTGCCTCTCG	7079
1141	Qy	AACACAACCTTAGATTCOAACAGAAAATCCCAACAAGGACCAAGACACCAACAGCAACACC	1200
7080	Db	ATAACAACCTCGCTCCAAAGGGGAAATCCCAACAAGCAAAAGAGCACACGACAGCAAAAAA	7139
1201	Qy	AACATCATCATGACGACATCAGATATAAACAGCAAAACACCCCAACAAATCTTCTCCGGAT	1260
7140	Db	GGCCCCGCCAACACGGCACCAAAACACACAATAAGAGCATTTTCACAGTCTCTCCCCCAACC	7199
1261	Qy	TCTAGTCCGACAAACCCGCCCTCTCTATATATCTTTAGAAAGAAAGCATCGATCTCTGGAGG	1320
7200	Db	CCCAGCTCGACTGCACAACATCTTGATATTTCAAGAAAGCGAGTATCTCTCGAGG	7259
1321	Qy	GAAGGCGACATGTTCCCTTTTCTGGATGGGTTAATAAATGCTCCAATGATTTTGACCCA	1380
7260	Db	GAAGCGACATGTTCCCTTTTCTGGATGGGTTAATAAATGCTCCAATGATTTTGACCCA	7319
1381	Qy	GTTCCAANTACAAAACAACTTTTGATGATCCTCTAGATTTCTGGTCCCTCGGCTGAGGAA	1440
7320	Db	GTTCCAANTACAAAACAACTTTTGATGATCCTCTAGATTTCTGGTCCCTCGGCTGAGGAA	7379
1441	Qy	GATCAACATGCCTCCCCCAATATTAGTTTAACTTTATCTTTATTTTCCCTAATAATAATAG	1500
7380	Db	GATCAACATGCCTCCCCCAATATTAGTTTAACTTTATCTTTATTTTCCCTAATAATAATAG	7439
1501	Qy	AACACTGCCTACTCTGGAGAAAATGAGAATGATTTGATCGAGATTTAAGAATTTGGAGC	1560
7440	Db	AACACTGCCTACTCTGGAGAAAATGAGAATGATTTGATCGAGATTTAAGAATTTGGAGC	7499
1561	Qy	GTTTCAGGAGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTWTGCGCCCTGGAAAT	1620
7500	Db	GTTTCAGGAGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTWTGCGCCCTGGAAAT	7559
1621	Qy	GAAGGACTTTACACTGCTGTTTTTAATTAATAAATCAAAAACAATTTGGTCTGCAGGTTGAGG	1680
7560	Db	GAAGGACTTTACACTGCTGTTTTTAATTAATAAATCAAAAACAATTTGGTCTGCAGGTTGAGG	7619
1681	Qy	CGTCTAGCCAATCAAACTGCCAAATCTTTGGAACTCTTATTGAGAGTCAACAACCTGAGGAA	1740
7620	Db	CGTCTAGCCAATCAAACTGCCAAATCTTTGGAACTCTTATTGAGAGTCAACAACCTGAGGAA	7679
1741	Qy	AGAACATCTCCTTAATCAATAGACATGCTATTGACHTTCTACTCAAGATGGGAGGA	1800
7680	Db	AGAACATCTCCTTAATCAATAGACATGCTATTGACHTTCTACTCAAGATGGGAGGA	7739
1801	Qy	ACATGCAAAAGTCTTTGACCTGATTGTTGATCGGGATAGAAGACTTTGTCAAAAATATT	1860
7740	Db	ACATGCAAAAGTCTTTGACCTGATTGTTGATCGGGATAGAAGACTTTGTCAAAAATATT	7799
1861	Qy	TCAGAGCAAAATTGACCAAAATTAATAAAGGACGAAACAAAAGAGGGGACTGTTGGGGTCTG	1920

Db	7800	TCAGACCAAAATTGACCAAAATTA	AAAAAGACGACAA	AAAAAGAGAGGGGACTGGTGGGGCTCTG	7855
Qy	1921	GGTGGTAAATGGTGGACATCCGACTCGGGTGGTCTTACTAACTTTGGGCATCTTTTGGCTACTA	1980		
Db	7860	GGTGGTAAATGGTGGACATCCGACTCGGGTGGTCTTACTAACTTTGGGCATCTTTGCTACTA	7919		
Qy	1981	TTATCCATAGCTGCTCTGTATGCTCTATCTCTATCTCTGATCTCTGATCTCTTACTAAATATATC	2040		
Db	7920	TTATCCATAGCTGCTCTGTATGCTCTATCTCTGATCTCTGATCTCTTACTAAATATATC	7979		
Qy	2041	GGATAA	2046		
Db	7980	GGATAA	7985		
LOCUS	MVBPCYC 19104 bp DNA linear VRL 19-SEP-1996				
DEFINITION	Marburg virus genes for vp35, vp40, vp30, vp24, glycoprotein, nucleoprotein, polymerase.				
ACCESSION	212132 S55429				
VERSION	212132.1 GI:541780				
KEYWORDS	glycoprotein; nucleoprotein; polymerase; vp24 protein; vp30 protein; vp35 protein; vp40 protein.				
SOURCE	Lake Victoria marburgvirus				
ORGANISM	Lake Victoria marburgvirus				
REFERENCE	Viruses; sgRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.				
AUTHORS	1 Sanchez,A., Kiley,M.P., Klenk,H.D. and Feldmann,H.				
TITLE	Sequence analysis of the Marburg virus nucleoprotein gene: comparison to Ebola virus and other non-segmented negative-strand RNA viruses				
JOURNAL	J. Gen. Virol. 73 (Pt 2), 347-357 (1992)				
MEDLINE	92166742				
PUBMED	1538192				
REFERENCE	2				
AUTHORS	Muhlberger,E., Sanchez,A., Randolph,A., Will,C., Kiley,M.P., Klenk,H.D. and Feldmann,H.				
TITLE	The nucleotide sequence of the L gene of Marburg virus, a filovirus: homologies with paramyxoviruses and rhabdoviruses				
JOURNAL	Virology 187 (2), 534-547 (1992)				
MEDLINE	92188528				
PUBMED	1546452				
REFERENCE	3				
AUTHORS	Feldmann,H., Muhlberger,E., Randolph,A., Will,C., Kiley,M.P., Sanchez,A. and Klenk,H.D.				
TITLE	Marburg virus, a filovirus: messenger RNAs, gene order, and regulatory elements of the replication cycle				
JOURNAL	Virus Res. 24 (1), 1-19 (1992)				
MEDLINE	92327834				
PUBMED	1626422				
REMARK	(sites)				
REFERENCE	4 (bases 1 to 19104)				
AUTHORS	Will,C., Muhlberger,E., Linder,D., Slenczka,W., Klenk,H.D. and Feldmann,H.				
TITLE	Marburg virus gene 4 encodes the virion membrane protein, a type I transmembrane glycoprotein				
JOURNAL	J. Virol. 67 (3), 1203-1210 (1993)				
MEDLINE	93172334				
PUBMED	8437211				
REFERENCE	5				
AUTHORS	Feldmann,H.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-MAY-1992) Feldmann H., Zentrum fuer Hygiene der Philipps-Universitaet, Institut fuer Virologie, Robert-Koch-Str. 17, Marburg, Germany, 3550				
REMARK	revised by [5] MAT				
REFERENCE	6 (bases 1 to 19104)				
AUTHORS	Feldmann,H.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-SEP-1994) Feldmann H., Zentrum fuer Hygiene der Philipps-Universitaet, Institut fuer Virologie, Robert-Koch-Str.				

17, Marburg, Germany, 3550
On Sep 22, 1994 this sequence version replaced gi:60623.
Location/Qualifiers
1..19104
/organism="Lake Victoria marburgvirus"
/mol_type="genomic DNA"
/strain="Musoke"
/db_xref="taxon:11269"
47..2843
/gene="np"
47..2843
/gene="np"
/standard_name="gene one"
/citation=[1]
/citation=[3]
/evidence=experimental
104..2182
/gene="np"
/standard_name="np"
/citation=[1]
/citation=[3]
/codon_start=1
/evidence=experimental
/product="nucleoprotein"
/protein_id="CAA78114.1"
/db_xref="GI:60624"
/db_xref="GOA:P27588"
/db_xref="UniProt/Swiss-Prot:P27588"
/translation="MDLSLLELGTPTAPHRNKKVILFDTNHQVSIQCNIIDAINS
GIDGLLEGLLTCVEHYNSDKDENTSPVAKYLRDAGVEFDVIKNADATPLDV
SPNEPHVSPLLALKTLESTESQGRIGLFLSFCSLFPLKLVGDRASIEKALQVTV
HQEGIVTPHNHLTGHMKVIFGILRSSFLKFLVLIHQGVNLVTGHDAYSILISNSV
QATRFSGLLIVKTVLEFLQKTDSDGTVLHLPLVTSKVNVSFKQALNSLRHGEYA
PFARVLNSGNLEHGLYQLSAIALGVATAHGSTLAGVNVGQYQQLREAAHDAEY
KLQRHEHQETQAIAEDDEERKILEQPHLQKTEITHSOTLAVLSQKREKLARLAAEIE
NNIVEDQGFKOSQNVFSFLNDPTPVETVQVAPRMPNRPALTLPVDDKIHESTEDS
SSSSFDLNDPFFALLNEDETLDDSVIMPCTTSREFGIPEPRQSDLNNSQKQE
DESTRIKQFLRQELPPVDEDESEYITDSQSDIQGSDNSQGVLDLPPPPYLAQE
KRQDPIQHPAANQDQPFQSDGVNGDILEPIRSPSPAPQEDTRMEAYELSPDFTN
MSFDEQMDLR"
2851..4409
/gene="vp35"
2851..4409
/gene="vp35"
/standard_name="gene two"
/citation=[3]
/evidence=experimental
2944..3933
/gene="vp35"
/standard_name="vp35"
/citation=[3]
/codon_start=1
/evidence=experimental
/product="vp35"
/protein_id="CAA78115.1"
/db_xref="GI:60625"
/db_xref="UniProt/Swiss-Prot:P35259"
/translation="MWDSSYMQVSEGLMTGKVPIDQVFGANPLEKLYKRRPKGTGVC
LQSCPLMSKATSDDIWDQLIVKRTLADLLIPINROI SDIQSTLSVTRVHEIER
QLHEITPVKMGRTLEALSKGSEMIAKYDHLVISTGRTAPAAAFDAYLNEHGVPPP
QPAIPKDIGVQAQKSGKGTWKNATDAAKMSKVLSESETFSKPNLSAKDLALLLF
THLFQNNTPPHILAQVLSIAKYSGKSGAFDAFHQILSEGENAQAALTRLSRTFDFAF
LGVVPVTRVKNFTVTPRPSQSLRAVPNPNTIDKGWVCVYSSEQGETRALKI"
4412..5818
/gene="vp40"
4412..5818
/gene="vp40"
/standard_name="gene three"
/citation=[3]
/evidence=experimental
4567..5478
/gene="vp40"

gene

mRNA

CDS

/standard_name="vp40"
/citation=[3]
/codon_start=1
/evidence=experimental
/product="vp40"
/protein_id="CAA78116.1"
/db_xref="GI:11552357"
/db_xref="GOA:P35260"
/db_xref="UniProt/Swiss-Prot:P35260"
/translation="MASSSNYTYMYLNSPPYADHGANQLIPADQLSNQOQITPNYV
GDLNLDQFKNGVCHAFTEALIDISAYNERTVKGVPALWPLGIMSNPEYPLAHTVAA
LLTGSYTIOTFTHNGQKQFVRNRLGTGIPAHPLRMLREGNOAFIQNMVIPNFSFNQF
TYNLTNLVLSVOKLPDDAMPKDKLIGNTMHPAVSIHPNLPPILVLPVKQAQROHK
NPNGPPLAISGILHQLRVEKVPKTSLSLFRISLPADMFSVKEGMMKRGKSPVYFYFQ
APENFPLNGFNRRQVVLAYANPTLSAV"
5822..8666
/gene="gp"
5822..8666
/gene="gp"
/standard_name="gene four"
/citation=[3]
/evidence=experimental
5940..7985
/gene="gp"
/standard_name="gp"
/citation=[3]
/codon_start=1
/evidence=experimental
/product="glycoprotein"
/protein_id="CAA78117.1"
/db_xref="GI:60627"
/db_xref="GOA:P35253"
/db_xref="UniProt/Swiss-Prot:P35253"
/translation="MKTTCFLISLILQGTKNLPILFIASNNQPNQVDSVCSGLTQKT
EDVHLMGFTLSGQKVDSPLEASRWAFRTPGVPKNVEYTEGEBAKTCYNISVTDPSG
KSLLLDPTNTRIDVPKCTTHIIQGNPHAGI ALHLWGAFLVDRIASTMYRGKVF
TEGNIAMIVNKTVHKMIFSRQGGYRHMNLTSNKYWTSSNGTQNTDTCGCGALQEY
NSHTQKCAKSPPPPLPTARPEIKLTSTPTDKLNTDPSDEEDLATSGSGSGER
EPHTSDAVTKQSLSTMPPTSPQSTPQGGNNTHSDQAVTELDKNNTAQPSMP
PHNTTSTNNISKHNFSTLSAQLQNTNTDNTQSTITENEQTSAPITLPTGNTPT
AKSTSKKGATPATNTNEHTSPPTPPSTAQHLVYFRKRISILWREGMFPFLDG
LINAIDPDVPNTKTI FDESSSGASAEEDQHASPNISLTSYPMINENTAYSGEN
ENDCABLRIVSVEQDDLAAGLSWIPFGGIEGLYTAVLIKGNLNLVCLRLRANQT
AKSLELLRLVTEERTFSLNRHAIDFLTRWGCTCKVLGPDCCIGIEDLSKNISEQI
DQIKKQKQEGTGWLGKRWMTSDWGLVTNLGILLLSIAVLIALSCICRTFTKYIG"
8762..10009
/gene="vp30"
8762..10009
/gene="vp30"
/standard_name="gene five"
/citation=[3]
/evidence=experimental
8864..9697
/gene="vp30"
/standard_name="vp30"
/citation=[3]
/codon_start=1
/evidence=experimental
/product="vp30"
/protein_id="CAA78118.1"
/db_xref="GI:541782"
/db_xref="UniProt/Swiss-Prot:P35258"
/translation="MQQPRGRSRTRNHQVPTIYHETQLPSKPHYTYNHPARMSST
RSSAESPTNHIIPARPSTFNLSKPPPPKDMCRNMKIGLPCADPTCNRDHLDNLT
NRELLLMARKMLPNTDXTFRMPQDCSPSLSKGLSKDKQKQTKDVLTLNGLHLSY
LHRSEIGNWRHRLRAULSLTCAGIRKTRNSLINTMLHMHENLPPQNGVYIKQYTT
GIHLDRGGQFEALMQGDKRSISLFFVQAALYVMNNIPCESSISVQASYESFYSSKS
R"
9990..11278
/gene="vp24"
/gene="vp24"

gene

CDS

Query Match

74.0%; Score 1514.8; DB 14; Length 19104;

Best Local Similarity 83.8%; Pred. No. 0; Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;									
Qy	1	ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT	60						
Ds	5940	ATGAAGACCATATGTTTCTCTATAGTCTTATCTTAAATCAAGGACAAAATCTCC	5999						
Qy	61	GTTTTGAATAATGCTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTCTCGGAACC	120						
Ds	6000	ATTTTGAAGATAGCTAGTAATAATCAACCCAAATGTGGATTGGTATGCTCCGGAAT	6059						
Qy	121	CTCCAAAGACAGAAGATGTTTCATCTGATGGATTTTACATGAGTGGGCAAAAGTTGCT	180						
Ds	6060	CTCCAAAGACAGAAGACGTCCATCTGATGGGATTTACACTGAGTGGGCAAAAGTTGCT	6119						
Qy	181	GATTCCTCTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC	240						
Ds	6120	GATTCCTCTTGGAGCATCCAGGATGGGCTTTTCAGGACAGGTGTTACCTCCCAAGAAC	6179						
Qy	241	GTGAGTATACGGAAGGAGAGAACCAAAACATGTTACAATATAGTGTAAACAGACCT	300						
Ds	6180	GTGAGTATACAGAGGGGAGGAGCAAAACATGTTACAATATAGTGTAAACAGGATCC	6239						
Qy	301	TCTGGAATATCTTCTGCTGTGATCTCCAGTAATATCCGGGATATCCCTAAATGTAAA	360						
Ds	6240	TCTGGAATATCTTCTGCTGTGATCTCTCAACATCCGCTGACTATCTTAAATGTCAA	6299						
Qy	361	ACTGTTTCATATATCAAGGTCAAAACCTCATGACAGGGGATGCTCCATTTGTTGG	420						
Ds	6300	ACTATCCATATATCAAGGTCAAAACCTCATGACAGGGGATGCTCCATTTATGTTGG	6359						
Qy	421	GGGGCATTTTCTGTATGATCGCTTGCTCTACAACAATGTACCGAGGCAAGTCTTC	480						
Ds	6360	GGAGCATTTTCTGTATGATCGCATTGCTTCCACAACATGTACCGAGGCAAGTCTTC	6419						
Qy	481	ACTGAAGGAATATAGCAGCTATGATGTTTAAAGACAGATTCACAGAAATGATTTTCT	540						
Ds	6420	ACTGAAGGGAACATAGCAGCTATGATGTTTCAATAAGACAGATGCACAAAATGATTTCTCG	6479						
Qy	541	AGGCAAGGACAGGTTATCGTACATGAATCTGACCTCCACCAATAAATATTTGGACAGC	600						
Ds	6480	CGGCAAGGACAGGTTATCGTACATGAATCTGACTTCTACTAATAAATATTTGGACAAGT	6539						
Qy	601	AGCAATGAACACGAGAAATGATCGGGATGTTTGGCATCTCCAAAGAAATACAATCC	660						
Ds	6540	AGTAACGGNACGCAACGNAATGACATGGATGTTTCGGGCTCTTCAAGAAATACAAATCT	6599						
Qy	661	ACAAACAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAATCCG	720						
Ds	6600	ACAAAGAACCAAAATGCTGCTCCGTCCAAATAATCCCTCCACCACTGCCACAGCCGCTCG	6659						
Qy	721	AGCATTCATCTACAAATACTCAAAATTAATCTGCTAAATCTGGAATCTAGAACCAAGT	780						
Ds	6660	GAGATCAAACTCACAGACACCCCAACTGATGCCAACCAATACACCGGACCCCAAGC	6719						
Qy	781	AGCGAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCACACACA	840						
Ds	6720	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA	6779						
Qy	841	ACTCTTAATGATGACCTGGAACAGAAACAAATCGTCAACAATATTTGTCACCTCTTCACTA	900						
Ds	6780	ACTTCTGATGCGGTCAACAGCAAGGGCTTTTATCAACAATGTCACCACTCCCTCACCAC	6839						
Qy	901	CATCCAAAGCCTCACAACTGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAAT	960						
Ds	6840	CAACCAAGCAGCCACAGAGGAGAAACAACAACCACTTTCCCAAGATGCTGTGACT	6899						
Qy	961	GAGCAATGGAACCGACCCCAACACACACAGCAACGCTCCTCAACAATACTAATACA	1020						
Ds	6900	GAACTAGACAAATAACACAACTGCAACACGCTCCATGCGCCCTCATTAACACTACCAACA	6959						
Qy	1021	ACTCCACCTTAAACATCTCAAGTACAACTGACTCTCTTCTCCCTCCACACCGCAAC	1080						

LOCUS AY430366 19112 bp RNA linear VRL 06-FEB-2004
 DEFINITION Lake Victoria marburgvirus strain pp4 guinea pig nonlethal variant, complete genome.
 ACCESSION AY430366
 VERSION AY430366.1 GI:40388379
 KEYWORDS
 SOURCE
 ORGANISM Lake Victoria marburgvirus
 Lake Victoria marburgvirus
 Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.
 REFERENCE 1 (bases 1 to 19112)
 AUTHORS Hevey, M., Negley, D., Geisbert, J., Jahrling, P. and Schmaljohn, A.
 TITLE Antigenicity and vaccine potential of Marburg virus glycoprotein expressed by baculovirus recombinants
 JOURNAL Virology 239 (1), 206-216 (1997)
 MEDLINE 98087840
 PUBMED 9426460
 REFERENCE 2 (bases 1 to 19112)
 AUTHORS Chain, P.S.G., Malfatti, S.A., Hajjaj, A., Vergez, L.M., Do, L.H., Smith, K.L. and McCreedy, P.M.
 TITLE Direct Submission
 JOURNAL Submitted (09-OCT-2003) Viral Sequencing Group, Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA
 REFERENCE 3 (bases 1 to 19112)
 AUTHORS Ichou, M.A., Paragas, J., Jahrling, P.B., Ibrahim, M.S., Loftis, L., Hevey, M. and Schmaljohn, A.
 TITLE Direct Submission
 JOURNAL Submitted (09-OCT-2003) Viral Genomics Group, USAMRIID, Fort Detrick, MD 21702, USA
 FEATURES
 source
 1. 19112
 /organism="Lake Victoria marburgvirus"
 /mol_type="genomic RNA"
 /strain="pp4 guinea pig nonlethal variant"
 /isolate="Musoke Kenya, 1980"
 /db_xref="taxon:11269"
 104..2191
 /gene="NP"
 104..2191
 /gene="NP"
 /note="nucleoprotein"
 /codon_start=1
 /product="NP protein"
 /protein_id="AAR85453.1"
 /db_xref="GI:40388380"
 /translation="MDLHSLLELGTKEPTAPVHRNKKVILFDTHQVSCINQIIDAINS
 GIDGLLEGGLTLCVHEHYNSDKDFNTSPIAKYLRDAGYFDVVKNAADATRFIDV
 IPNPHYSPLILALTKLESTESQGRIGLFLSCFLPKLVGDRASIEKALROTVV
 HQEGIVTYPNHWLTGGMKVI FGLRSSFLKFLVHQVNLVGTGDAVDSIISNV
 GQTRFGLIVKTVLEFILQKTDGVTLHPLRTSKVKNVASFQALSLARHEYA
 PFARVLNSGINNLEHGLYPQLSALGLAVATAGSTLAGVNVGEGYQQLREAAHDAEV
 KLQRHEHQIQAIEADEERKILEQFLQKTEITHSQTLAVLSQKEKLAAAEIE
 NNIVEDQGFQSONRVSQSLNDPTVEVTQAPRMNRTALPPVDKTEHSTEDS
 SSSSSFDLNDPFAALLNEDETLDDSVMI PCTTSREFOGI PEPPROSQDLNNSQKOE
 DESTNPILKQFLRQELPQVEDSEYITDSQESIDQPSDNEQGVLDLPPPLPQAQ
 KRQDPIQHPANQPPGSGIDVNGDILSPRSSPSAQEDCKMREAYELSPDFTN
 DEDQNNQPPQVTVTKGRTFLYNDLLQTNPPESLITALVEEQNPVSAKELQADWPD
 MSFDERHRVAMNL"
 2945..3934
 /gene="VP35"
 2945..3934
 /gene="VP35"
 /note="minor component of RNA-dependent RNA-polymerase"
 /codon_start=1
 /product="VP35 protein"
 /protein_id="AAR85454.1"
 /db_xref="GI:40388381"
 /translation="MWDSSYMQVSEGLMTGKVPIDQVFGANPLKLYKRRKPKGTG
 LQCSFLMSKATSDTDIIWDLIVKRTLADLLIPINQISDIQSTLSEVTVRHEIER
 QLHEITPVLKMRLEAISKGMSEMLAKYDHLVISTGRTTAPAAFDAYLNEHGVPVP
 QPAIFKDLGVAQAQCSKGTWKNATTDAADKMSKVLSETFSPKNSAKDLALLIF
 TLPFGNTPPHILAQVLSKAYKSGKSGAFDLDFHQLISEGNAQAALTRLSRTDFAF
 LGVPPVIRVKNFQTVPRPCQKSURAVPPNPITDKGWCVSYSEQQGGETRAIKI"
 4568..5479
 /gene="VP40"
 4568..5479
 /gene="VP40"
 /note="matrix protein"
 /codon_start=1
 /product="VP40 protein"
 /protein_id="AAR85455.1"
 /db_xref="GI:40388382"
 /translation="MASSSNNTYMQVLYNPPPPVADHGANQLIPADQLSNQOQITPNVY
 GDNLNDQFGKNVCHAFTELEAIIIDISAYNERTVKGVPAMVPLGITMSNFEPYLAHTVAA
 LITGSYTIQFTHNGQKFRVNRNLGTGIPAHPLRMREGNOAFTQNNVIFRNFSTNOF
 TYNLTNLVSVOKLPDDAWRPSKDKLIGNTMHPAVSIHPNLPTVLPTVKKQAVROHK
 PNPNGLLALSGILHQLRVEKVPKTSFLFISLPADMFSVKSGMKRGENSPVYFQ
 APENFPLGNFRNQVVLAYANPLSAV"
 5941..7986
 /gene="GP"
 5941..7986
 /gene="GP"
 /note="envelope transmembrane glycoprotein"
 /codon_start=1
 /product="GP protein"
 /protein_id="AAR85456.1"
 /db_xref="GI:40388383"
 /translation="MKTTCFLISLLIQTENLPILLEIASNNQPNQVDSVCSGTLQKT
 EDVLMGFLISGQKVADSPLEASRWAFTGVPVKVEYTEGEAEKTCINISVDPDSG
 KSLLDPPNIRDYPKCKTIHHIQGNPHQAQIALHLWGAFFLYDRIASTVMYRGVVF
 TEGNIAAMIVNKTVMIFSRQGGYRHNLTSTNKYTSNGTQTDNLTCFGLAQSY
 NSTKNQTCAPSKIPPLPTARPEIKLTSTDTATKLTNTSPSSDDEDLATSGSGSER
 EPHTSDAVTKGLSSMTPTSPQSPSTPOGGNNTNHSODAVTELDKNTTAQPSMP
 PHNTTISNTNTSKHNPSTLSAPLQNTDNTQSTITENQTSAPSTITLPPGNPTT
 AKSTSKGPAITAPNTNEHFTSPPTPSPTAQHLVYPRKRSILMREGMDPPFLDG
 LINAPIDFDPVNTKTI FDESSSGASAEEDHSPNISLTLSYFPNINENTAYSGEN
 ENDDAELRIWVQEDDLAAGLSWIPFGGIEGLYTAGLIKQNNLVLCRLANQAT
 AKSLELLRVTTERTFSLINRHAIIDPLLRWGTCKVGPDCGICIGEDLSKNISEOI
 DOJKDBOKEGTGWLGGKWTSDWGLTVLNLGILLLSIAVLIALSCIRFTKYIG"
 8869..9714
 /gene="VP30"
 8869..9714
 /gene="VP30"
 /codon_start=1
 /product="VP30 protein"
 /protein_id="AAR85457.1"
 /db_xref="GI:40388384"
 /translation="MQQPRGRSTRNHQVPTTYHETQLSPSKPHYNYHPRARMSST
 RSSAESPTNHI PRARPPSTFNLSKPPPKDMCRNNKIGLPCADPTCNDRHDIDNLT
 NRLLLLMARKMLPNTDKTFRSPQDCGSLSKGSKDKOEQTKDVLTLNGLHILSY
 LHRSEIGKDETSLSRAALSITCAGIARTNRESLINTWELMHNHNL PQDQNGVTKQY
 TGHLDKGGQFEALWQGDWRKRSISL FVQAALYMNIPCESSISVQASVDHFILOS
 QGKGQ"
 10207..10968
 /gene="VP24"
 10207..10968
 /gene="VP24"
 /codon_start=1
 /product="VP24 protein"
 /protein_id="AAR85458.1"
 /db_xref="GI:40388385"
 /translation="MAELSTRYNLPAVNTENSINLDINSTARWIKEPSVSGWTVKQGN
 FVPHIYNTGNTLLHLKSNFVPEWQOTRNLFSHLKPKSTIEPPLAIRILLGVAL
 KDQELQOLIPGRSIVHMLSEWILLEVTSAIHSPLNLGIYLTSMFKLMAGVRNF
 FNMFTLHVVDNHGKPSSEIKLTGQOIITIRVMGFLVEVRIDIBPCCGETVLSBS
 VFWGLVAEAVLREHSGMEKQQLNLQTYMNSKIAI"
 11481..18476
 /gene="L"
 11481..18476
 /gene="L"
 /note="major component of RNA-dependent RNA-polymerase"
 /codon_start=1
 /product="L protein"
 /protein_id="AAR85459.1"
 /db_xref="GI:40388386"
 /translation="MQHTQYDARIUSSPILQDCLLARSGLYSHYSHNPKNRNCR

1PHHYLRNSTALTKFLQNCISILTVPPHPSIWDHILTSIQYDAINHVDVDFKYLLPSEL
VKYANWDEFLKAYKIGTDLGLDGHVFSASRQSCDFSPKENPYWYMGMLLVHLSQLAR
RIKQGNDFLRNWKYIGTDLGLFIADPFIKVPVKTKIRNVAVSLOASQKGLRIWYRD
QNLTPYLCDDREFIVSAYVECFIMIKDVFIERVNTWEICARAWLESDPGADYPLDVL
GELYNQGOQIITAMYLEDSQFKLKHLEPLCVSCIQTHGIFTPRKXWFQSQMIKSYDEL
HDNLKLGQISOMKAECAQNFIKTIIVQAKLTPOQYCELSFLOKHGHPVLYINDVALDKV
KKHAQSTIKLKPKNMFETFCVFKPIKAKXHYHSQSWYKTHDHLHTPLRQHVSNVS
FPQAEIYQHLWMEYFVEHEPLFSTKIIISDLSIFIKORATVANOECMDSVDFRSVLGY
NPLVRFQSKRVPEQLOALDGLAKFPPSNMVMVTEREOKALLHSHWHNSASIGEN
GKLPVVRNVOTLAEALLADGLAFPPSNMVMVTEREOKALLHSHWHNSASIGEN
AIVRGASFVTDLEKYNLAFRYEPFRHFDYCNRCYGVKNLFDWVHFLIPLCYMVUSD
YSPHCIVTDRNRNPPDCANAYHHLGEGLEQKQKWLTCISCAQITIVELKTKUKLKS
SVMGDQCITLTFSPDIPNDYQENEAELNAAVAVELAITTYGSGIFLKPBEETVH
SGFIYFGKQVINGQLPSQNTWARKCPLSDISDFDQLGSLASIGTSFERTGETRH
IFPSRWIASFMSLAINLNQNLHGFPLGFNIDICPKPKLTFSEKLIJALITPVOVLG
LFLNPEKLFVRNISDPLTSGLFOLKNALEFLEKEELFYILISKKPGLADASDFVMP
LGLNVPGSKKEITFLRQTVRENIITTSQNRINLSLPHIGSDLEORVCEWILLSNPVM
SRPAADIFSRPSGRKQLVGLGERTILLASRTISLTTGEMLMKURKRELTRNRKWS
FSYDALDDLSLESLEKFTCTVDVANFLRAYWSVDLKGKRLIGATPLCLLEQFEVX
INLSDBREQNLSDSKSTINLIPYDCCKLEKLEGSNDTELNYVSCALDRKVQOKHPS
VNRLAWTIGNRAPYIGSRTEDEKIGYPPLRVNCPSAALKEATEMVSRLILWVTQADRE
KLILPLANSRVNLDYQTVNLPLPHYSGNIVHRVNDXGOHFSWANRMSNTSTRAILS
TNTLGKXAGGQAAIDSNIIIFONTINLGVAVLDIALSLAKLSASNVTFLRLMKKCTT
RHPSEXIFYDKPLDVLNKNYMDNELVYDNDPLSGIKRGRVRSRSTLSLSLVSDI
GSYDPTTAATLGTETIVGSIISDESQSDTPISSGCTKTFVTHFLVTPVESIFVAFG
ANLIVESLSLRKISIKNLSDLTFLISSTIRNLSRILIKQSTFRHELVLTRLAHHI
PLISIMLGSAGEKSSDAVRLFLTASVONFINNFSCIMKKGQSSLPVLPVPSGEOQ
LKPIILKILQRLSDLLSPKIKRKLIAOTCCPIGSPFWYVPSKSTRNHYASLNYWRD
KANKVNTFFSHLNCISPEFSSTSSVSNQOVNSKIIVIPENITBEINAKTRLNY
GSTALQMDMTKQSEQLNVENRPEGIRFKNQKIKTHDORERESPOOMPED
NMQTPAHTHSSPFOILKISDAHEDFASKIILNSENINLNLTTEYTLNKLTLTPTR
TEILDTSPQRSYSTGRKSLRSREQSYLYVDCSNIPGISLDPGPRSMSONOVO
MLINTYKRDLHACPSNDFCRFTGVVSSMHVKLYDLPGLPKKAICLAEGEGSARL
LLKWKETDYLFPNTLATSQEAELISGRVIPRLMYNLDRLSLESRRLLINLITQ
ITDITNPLWLSOYLPEDSDILTMDAETKDEOREQYIKIIVINWTRSPNPKIS
IIKVELLYDEGTFRLMKNAIQQYGVQKPKXPYSNAKSEWYLCGKRRIORLOQDFS
DOVGFLICKAMSRQRAIPYWLKHIEKNYPASLEFPLTLGFPSELSFCHRYTIFP
SEGKALFHVKVOSYVROQHLSLMLDVENNSPLDLRNFHICSLRGKIKTYKYNIDLK
LNILVTI KAVEKGNWSOLVELLIPNHSVCIVVHDHSCSCEKRLLLKLDPIRNTKIAEQ
KLLNRVIGYILPPFPGLFKSGSLRA"

ORIGIN

Query Match 74.0%; Score 1513.2; DB 14; Length 19112;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 1713; Conservative 0; Mismatches 333; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATATTTCTGATTAGTCTCATATTTAAATCCAAAGTATATAAACTCTCCCT 60
DB 5941 ATGAAGACCATATGTTTCTTATCATGCTTATCTTTAAATTCAGGGACAAAAAATCTCCC 6000
QY 61 GTTTTAGAAATGTCTAGTAACGCCAACCTCAAGATGTAGATTCAAGTGTGCTCCGGAACC 120
DB 6001 ATTTTAGAGATAGCTAGTAAATAATCAACCCCAAAATGTGGATGCTGCTCCGGAAT 6060
QY 121 CTCCTCAAGACAGAGAAGATGTTTCATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT 180
DB 6061 CTCCTCAAGACAGAGAAGCTCCATCTGATGGATTCACACATGAGTGGGCAAAAAGTTGCT 6120
QY 181 GATTCCTCTTTCGAAGCATCTAAAAGATGGGCTTTTCAGACAGAGTGTCTCTCCCAAGAAC 240
DB 6121 GATTCCTCTTTCGAAGCATCTCAAGGATGGGCTTTTCAGACAGAGTGTACCTCCCAAGAA 6180
QY 241 GTTGAGTATACGAGAGGAGAGCAAGCAAAACATGTTACATATAGGTGTAACAGACCT 300
DB 6181 GTTGAGTATACGAGAGGAGGAGAGCAAAACATGCTACATATAGGTGTAACAGGATCCC 6240
QY 301 TCTGGAAAATCTCTGCTGCTGATCCTCCAGTAATATATCCGGGATTTACCTAAATGTATAA 360
DB 6241 TCTGGAAAATCTCTGCTGCTGATCCTCTACCAACATCCCGTACTATCCTTAATATGCAA 6300
QY 361 ACTGTTATCATATTTCAAGGTCAAAACCTCTATGACAGAGGATTCCTCATTTGTGG 420
DB 6301 ACTATCCATCATATTTCAAGGTCAAAACCTCTATGACAGAGGATTCCTCATTTATGG 6360

```
QY 1501 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGTGATGCAGAGTTAAGAATTTGGAGC 1560
Db 7441 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGTGATGCAGAGTTAAGAATTTGGAGC 7500
QY 1561 GTTCAGGAGGATGACTCTGGCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTTGGAAATT 1620
Db 7501 GTTCAGGAGGATGACTCTGGCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTTGGAAATT 7560
QY 1621 GAAGGACCTTTACACCTGCTGTTTTTAATTAATAATCAAAACAATTTGGTCTGCAGGTTGAGG 1680
Db 7561 GAAGGACCTTTACACCTGCTGTTTTTAATTAATAATCAAAACAATTTGGTCTGCAGGTTGAGG 7620
QY 1681 CGTCTAGCCAATCAAACTGCCAATCCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
Db 7621 CGTCTAGCCAATCAAACTGCCAATCCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 7680
QY 1741 AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTACTACAAAGATGGGAGGA 1800
Db 7681 AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTACTACAAAGATGGGAGGA 7740
QY 1801 ACATGCAAAAGTCTTGGACCTGATTGTTGCAATCGGGATAGAGACTTGTCCAAAAATATT 1860
Db 7741 ACATGCAAAAGTCTTGGACCTGATTGTTGCAATCGGGATAGAGACTTGTCCAAAAATATT 7800
QY 1861 TCAGAGCAAAATGACCAATTAATAAGACCAACAAAGAGGGGACTGTTGGGGCTG 1920
Db 7801 TCAGAGCAAAATGACCAATTAATAAGACCAACAAAGAGGGGACTGTTGGGGCTG 7860
QY 1921 GGTGTAATGCTGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTGCTACTA 1980
Db 7861 GGTGTAATGCTGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTGCTACTA 7920
QY 1981 TTATCCATAGCTGCTTGAATGCTATCTATCTGATTTGCTATCTTTTACTAAATATATC 2040
Db 7921 TTATCCATAGCTGCTTGAATGCTATCTGATTTGCTATCTTTTACTAAATATATC 7980
QY 2041 GGATAA 2046
Db 7981 GGATAA 7986

RESULT 7
AY430365 19113 bp RNA linear VRL 06-FEB-2004
LOCUS Lake Victoria marburgvirus strain pp3 guinea pig lethal variant,
DEFINITION complete genome.
ACCESSION AY430365
VERSION AY430365.1 GI:40388387
KEYWORDS
SOURCE Lake Victoria marburgvirus
ORGANISM Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 19113)
AUTHORS Hevey,M., Negley,D., Geisbert,J., Jahrling,P. and Schmaljohn,A.
TITLE Antigenicity and vaccine potential of Marburg virus glycoprotein
expressed by baculovirus recombinants
JOURNAL Virology 239 (1), 206-216 (1997)
MEDLINE 98087840
PUBMED 9426460
REFERENCE 2 (bases 1 to 19113)
AUTHORS Chain,P.S.G., Malfatti,S.A., Hajjaj,A., Vergez,L.M., Do,L.H.,
Smith,K.L. and McCready,P.M.
TITLE Direct Submission
JOURNAL Submitted (09-OCT-2003) Viral Sequencing Group, Lawrence Livermore
NATIONAL National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA
REFERENCE 3 (bases 1 to 19113)
AUTHORS Ichou,M.A., Paragas,J., Jahrling,P.B., Ibrahim,M.S., Lofts,L.,
Hevey,M. and Schmaljohn,A.
TITLE Direct Submission
JOURNAL Submitted (09-OCT-2003) Viral Genomics Group, USAMRIID, Fort
Detrick, MD 21702, USA
FEATURES Location/Qualifiers
```

```
source
gene
CDS
gene
CDS
gene
CDS
gene
CDS
```

```
1. .19113
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="pp3 guinea pig lethal variant"
/isolate="Musoke Kenya, 1980"
/db_xref="taxon:11269"
104..2191
/gene="NP"
/notes="nucleoprotein"
/codon_start=1
/product="NP protein"
/protein_id="AAR85460.1"
/db_xref="GI:40388388"
/translation="MOLLSLELGTPTAPHVRNKKVILFDTHNQVSIQNIIDAINS
GIDGLDLEGGLTLCVEHYNSDKOKFNTSPIAKYLRDAGYEFDVTKNADATFLDV
IPNEPHYPLIALKLTLESTESQGRIGLFLCSLFLPKLVGDRASTIEKALRQVTV
HQEGIVTYPNHMLTTGHMKVIFGILLRSSFILKVLHOGVNLVTHGDVDSITNSV
QOTRPSGLLIVKTVLEFILKQTSQGVTLHLVRTSKVNEVASFQKALSILARGEYA
PFARVNLGGINLLEHGLYPLSAIALGVATAHGSTLAGVNVGQYQOLREAAHDAEV
KLQRHEHQBIOIAEDDEERKILEQPHLQKTEITHSQTILAVSKREKILAAAEIE
NNTVEDGFKQSNRVSQSFNDPTPEVTVQARPMNRPTALPPVDDKIEHSTEDS
SSSSFDVLDNDFALLNDEDDTLDDSVMI PGTTSRBFQGIPEPRQSQDLNNSQKOB
DESTNPKKQFLRYOELPPVOEDDESFTTDSOESIDQSGDNEQGVLDLPPPIYAOE
KRQDPIQHPANFQDPFGSGIDVNGDILLEPIRSPSSSAQEDTRMREAVELSDFTN
DEDNQNPQRVVTTKGRTFLYNDLLQTNPPESSLITALVEEYQNPVSAKELQADWPD
MSFDERRHVAMNL"
2945..3934
/gene="VP35"
/notes="minor component of RNA-dependent RNA-polymerase"
/codon_start=1
/product="VP35 protein"
/protein_id="AAR85461.1"
/db_xref="GI:40388389"
/translation="MMDSSYMQQVSEGLMTGKVPIDQVFGANPLKLYKRRKPKGTVG
LQCSPLMSKATSTDDIWDQILVKTLDLIPINQIISDIQSTLSEVTVTRVHEIER
QJHEITFVLKMGRTLEAISKSGSEMLAKYDHLVISTGRTTAPAAFDAYLNKHGVPFP
QALIFKDLGVAQQAQSGKTVMKNATTDAAKMSKVLSELSETFSKNLSLKDALLFL
THLPGNNTPFHILAOVLSKIAVYKSGSGAFDAPHOILSEGENAQAALTRLSTRFADF
LGVPVPVIRVKNFQTVPRPCQKSLRAVPPNPTIDKGWVCYSSSQEGETRALKI"
4568..5479
/gene="VP40"
/notes="matrix protein"
/codon_start=1
/product="VP40 protein"
/protein_id="AAR85462.1"
/db_xref="GI:40388390"
/translation="MASSSNVNTYMOYLNPPYADHGANOLIPADQLSNOQGITPNYV
GDNLDDQFKNGVCHAFTLEALIDISAYNERTVKGVPALPLGIMSNFEXEPLAHTVA
LLTGSYITQFTNGKQKFNRLCTGIPAPHLRMREGNQAFQINWVIFRSTNQF
TYNLTNVLVSQVLKPDADAPSKNKLIGNTHHPAVSIHPNLPVLPVTKRQAQRQHK
NPNNGPLAISGILHQLRVEKPEKTSLSRISLPADMFVSKGEMMKRGNSPVYFQ
APENFPLNGFNROVLAYANPTLSAV"
5941..7986
/gene="Gp"
/notes="envelope transmembrane glycoprotein"
/codon_start=1
/product="Gp protein"
/protein_id="AAR85463.1"
/db_xref="GI:40388391"
/translation="MKTKTFLSILILIOGTKNLPFILEASNNQPNQVDSVCSGTLQKT
EDVHLMGFTLSGQKVPLEASKRWAPRTGPPKNVEYTEGEAKTCYNISVDPDG
KSHLLDPTNIRDYPKCKTIHHIQONPHQAQIALHLWGAFFLDYRGLASTMYRQKVE
TEGNIAMIVNKTVMKIFSRQOGYRHMNLTSNKYWTSNGTQNTDNGCFGALOEY
NSTKNQCAPSKIPPLPTARPEIKLSTPDTATKLANTDPSSEDEDLATSGSGSGER
EPHTTSDAVTKQGLUSTMPPTPSPQPGQGNNTNHSQDAVTELDKNTAQPSPMP
```



```
PHNTTITSTNNTSKHNFSTLSAPLQNTTNDNTQSTITENBQTSAPSITTLPTGNPTT
AKSTSSKGPATTPANTTNEHTGPTPTPSAQHVLVFRKRSLILRMGDMFPFLDG
LINAPIDPDPVRIKTTIDESSSGASAEQHQASPNLSLTYSPFNINENTASGEN
ENDCAELRWSVOEDDLAAGLSMIPPPGIGIEGLYTAGLIGNONNLVCLRRRIANOT
AKSLILRVTERTFSLINRHALDFLLTWGCTCKVLGPDCCIGIGIDRSITKTIQ"
DQIKDEQEGTGWGLGKWMTSDMGVLTNLGILLLLSIAVLIALSCICRTIKYIG"
8869. .9714
/gene="VP30"
CDS
8869. .9714
/gene="VP30"
/codon_start=1
/product="Vp30 protein"
/protein_id="AAR8546.1"
/db_xref="GI:40388392"
/translation="MQQPRGRSRTNRHOVPTIYHETOLPSKPHYTNXPHRAGMSST
RSSAESSPTNHPRARPPSTFNLSKPPPPKDMCMRNMKIGLPCADPTCNDRDHDLDNLT
NRELLNLWARKMLNTDKTFRSPQDCGSPSLSKGLSKDQKQSTQKDLVLTQNLHLSY
LHRSIGKLSLRAULSLTCAGIRKTRSLNTWELHNMHNELPDQNGVIAKQTY
TGIHLDGQGFEEALWQMDKRSLSLFOAALYVMNNIPCESSISVQASDYHFILOQS
QGGQ"
10207. .10968
/gene="VP24"
CDS
10207. .10968
/gene="VP24"
/codon_start=1
/product="Vp24 protein"
/protein_id="AAR8546.1"
/db_xref="GI:40388393"
/translation="MAELSTRYNLPANTVENSINLDNSTARWIKERSVSGWTVKGN
FVFHPNTGMILLHLKSNFVVPVQQTRNLFSLFNPKSTIIEPFLALRLLGVAL
KDQELQOQLIPFGRISVHMLSEWLLLEVTSAIHSPLNLGIYLSDMFKILMAGVRF
FNKFTLVHNDHGKSPSIEIKLQGOQIIITRVNMGFLVEVRRIDIEPCCGETVLS
VVFGLVABAVLRHSQMEKGQPLNLTYMNSKIAI"
11481. .18476
/gene="L"
CDS
11481. .18476
/gene="L"
/notes="major component of RNA-dependent RNA-polymerase"
/codon_start=1
/product="L protein"
/protein_id="AAR8546.1"
/db_xref="GI:40388394"
/translation="MQHPTQYDPARLSPPHIIQDCDLLARSGLYSHYSHNPKLRNCR
IPHLYLRNSTALKTFQNGSILTVPPHSIDHILTSIQYDAINHVDVDFKYLPSL
VKYAWDNELKYNLIGLDHVPASARQCEDFSPKENPYWGLMLLVHLSQLAR
RIKQWGSRLSNKWFIGTDLRGIADFVIKVPVKTIIRNAVSLQASKPGLRIYRD
QNLPYLCDDDEFIVSAVECFIMIKDVIERYNTWEICARAWLESDSGADYPLDVL
GELYNQGOQITAMYLEDDGFKLTKHLEPLVCVSCIQTHGIFTPRKYWFQSQMKSYYDEL
HDLNLQISDNKAECAQNFIKTIYQAKLTPOYCELSLQKWHGHPVLYNDVALDKV
KKHAQSTKILKPKVMFETFCVFKTIKXNHYSQGSWYKTTDLHLTPYLRQHIVSN
PFSQAEIYQHLWEYFVBEHPLFTKIISDLSIFIKDRATAVNOECWDSVDRSVLGY
NPPVRFQSKRVEQFLGQADFLSNQILSFSAEKLVLAPSFRNFSLSKEKELNIGRTF
GKLPYRVNVTQLABALLADGLAKAPPSNMVMVTEREQEALLHQASHHNSASIGEN
ALVRGASFVTDLEKTLNLAIFYETKHFIDYCNRCYGVKNLFDMMHFLIPLCYMHVSDF
YSPPHCVTEDNRNPPDCANAYHYHGLIEGLQKQLWTCISCAQITILVELTKUKLKS
CVMQNGQITITSLSPIDAPNDYQNEAEALNAARVAVELAITTGYSGIFLKPEPTFVH
SGFIYFGKQYVNLQPSLQKTMARCPGLSDSIDPDQGLASIGTSFPERGTSRTH
IFPSWIASFSMLAINLQNLHGFPIGFENIDISCPEKPLTFSEKILALIPTOVIGL
LSFLNPEKLFYRNISDPLTSLGFLQKNALEFLEKEELFYIILSKPGLADASDFWNP
LGLNVPFGKEIITFURQIVRENIITTSQNRILNSLFIHIGSDLEQVCEWLLSNPVN
SRFAADITSRTSPGVLGLEGTRTLASRTISLTTEGTMKLRLNTRNWKSW
FSYIDALDDLSLEKEFTCTVDVANFLRAYSWDLVKGRLIGATILPCLIEQFEVKW
INLSBDLQFNLSDSKSTINLLPYDCKELRLEGSNDTELNYVSCALDRKVVQKHS
VNRLAWTIGNRAPYIGSRTEDEKIGYPLRVNCPSAALKEAIEWVSRLLLWVTQGTADRE
KLLIFLLASRVNLDYQTVLNFELPHYSGNIVHRYNDQYQHSFMANRNSVTSTRAIS
TNLKGYAGGQAAIDSNIIIFQNTINLGVLDLIALSLKSSASNSVTFRMLNKCCT
RHVPSEXYLFQKPLVDNLKYNMDNELVDNPLCSGIGRIGRVSRLTSLSLNVSDI
GSYDPTTAAMTGLSTIVGSIIFSDESSQSTDPISGGCTKTFVTHFLVPVSEISYAFG
ANLIYVESLSGERSIKSNLSDLTFLISSTIRNLSRSLIIOQTFRHELVLTRAHHI
PLISLMGSGSAGEKSDPAVLFLTASYNFINNFSCIMKKGQSLPWLVPSEGOQ
LKPILKILQRLSDLLSPKIKRKLADTCCPIGFSFWYPSKSTRNTHYASLNYWRD
KANKVNTPFPSHLINCSFPFESSSTSVSSNOQVNSKIIVIPENITEINARTLLINY
GSTALQGMDTKTMPEQNLVENCPSREGIRFKNQNKIKTHDQRCEREBSPPQMFED
```

```

NMQTPAHIHSSSPQIILIKSLDAHEDFDASKIILNSENINLNTYTLNTKLLTTPTR
TEILDTPSPASSRYSSTRSRSLRSQASLYYDCSNIPSIDPFRSGMSDQNVQ
MLINTYKRDHLHAFDSNQFCFTQVSSMHYKLYDLLPPGKLKAICLAIEGSGSARL
LLKWKETDYLFNFTLATDSQEAELISGRVIRMLYNIIRLSALLESRRLLINLTIQ
ITDITNPLWLDSDVIQVLPEDSDILTMDAETTKDETRELQVKTIVNIWRTSPNIPKIS
IIKVLFDYEGTFLMKNAIQYGVQVQKPKPYSSNNAKSEWYLCGGRRIORLQIDFS
DQVGFILICKMSRQRAIPYWLKHIEKNYPASLHEFFLTILGFPFLSESSPHYTYTIFP
SEGKALFHVKQSVYRQOQHLHLSLMDYNNSPLLDLRNHFICSLRGKITKYNDILK
LNLVIKAVEKGNWSQVLEILPNMHSVCIVHVDHECSGCEKRLLLKDLFIRNTKIAEQ
KLLNRVIGYILFPFPGFLPKSGSLRA"
ORIGIN
Query Match 74.0%; Score 1513.2; DB 14; Length 19113;
Best Local Similarity 83.7%; Pred. No. 0;
Matches 1713; Conservative 0; Mismatches 333; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT 60
Ddb 5941 ATGAAGACCATATGTTTCCCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 6000
Qy 61 GTTTTATAGAAATTGCTAGTAACAGCCAACTCAAGATGTAGATTTCAGTGTGCTCCGGAAACC 120
Ddb 6001 ATTTTATAGATAGTAGTATTAATCAACCCCAAAATGTGGATTTCGTATGCTCCGGAAC 6060
Qy 121 CTCCAAAGACAGAAGATGTTTCATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT 180
Ddb 6061 CTCAGAAGACAGAAGACGTCATCTGATGGGATTTACACATGAGTGGGCAAAAAGTTGCT 6120
Qy 181 GATTCCTCTTTGGAGGATCTTAAACGATGGGCTTTCAGGACAGGTGTCCTCCCAAGAAC 240
Ddb 6121 GATTCCTCTTTGGAGGATCTTAAACGATGGGCTTTCAGGACAGGTGTCCTCCCAAGAAC 6180
Qy 241 GTTCAGTATACGGAAGAGAGAAGCCAAACATGTTTACAATATAAGTGTACAGACACCT 300
Ddb 6181 GTTCAGTATACAGAGAGGGAGGAAGCCAAACATGCTACATATAAGTGTACAGGATCC 6240
Qy 301 TCTCGAAAACTCTTGTCTGCTGGATCCTCCAGTAAATATCCCGGATATACCTTAAATGTAAA 360
Ddb 6241 TCTCGAAAACTCTTGTCTGCTGGATCCTTACCAACATCCCTGACTATCTTAAATGCAAA 6300
Qy 361 ACTGTTTCATCATATTCAGGTCAAAACCTCATGTCACAGGGGATTCCTCCATTTGCG 420
Ddb 6301 ACTATCCATCATATTCAGGTCAAAACCTCATGTCACAGGGGATTCCTCCATTTATGG 6360
Qy 421 GGGGCATTTTCTTGATGATCGCGTGTGCTCTACAACAATGTACCGAGGCAAGTCTTC 480
Ddb 6361 GGAGCATTTTCTTGATGATCGCATTCGCTCCACACAATGTACCGAGGCAAGTCTTC 6420
Qy 481 ACTGAAGAAATATAGCAGCTATGATTGTTAATAAGACAGTTCACAGAAATGATTTTTCT 540
Ddb 6421 ACTGAAGGAAACATAGCAGCTATGATTGTTCAATAAGACAGTTCACAAATGATTTCTCG 6480
Qy 541 AGGCAAGGACAGGTTATCGTCACATGACCTTGACCTCCACCAATATAATATTGGACAGC 600
Ddb 6481 CGGCAAGGACAGGAGGTACCGTCATAGATCTGACTTCTACTAATAAATATTGGCAAGT 6540
Qy 601 AGCAATGAAACCGCAGAGAAATGATACGGGATGTTTGGCATCTCTCCAGAAATATACAACTCC 660
Ddb 6541 AGTAACGGAAGCAAAACGATGACATGGAATGTTTGGCGCTCTTCAAGATACAAATCT 6600
Qy 661 ACAAAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAAGTCCG 720
Ddb 6601 ACAAGAACCAAAACATGCTCCGTCCGTCACAAAATACCTCCACCACTGCCCCACGCGCG 6660
Qy 721 AGCATTCACTTACAATATCTCAATTAATCTGTAATCTGGAACCTATGAACTTCAACCAAGT 780
Ddb 6661 GAGATCAAACTCAAGACACCCCACTGATGCCCAAACTCAATACCCAGGACCAAGC 6720
Qy 781 AGGCACGATAGGACCTTATGATTTTCGGGCTCAGGATCTGAGAAACAGGGGCCCCACACA 840
Ddb 6721 AGTGATGATAGGACCTTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 6780
Qy 841 ACTCTTAATGTAGTACTGAAAGAAACAAATCGTCAACAATATTGTTCCACTCTTCACTA 900
```

Db 6781 ACTTCGTGATGGGTCACCAAGCAGGGCTTTTCATCAACAATGCGACCCACTCCCTCACCA 6840
Qy CATCCAAGCACCTCACAACATGAGCAAAAGTAGCAATCTCTCCGACATGCTGTAACT 960
Db 6841 CAACCAAGCAGCCACAGCAGGAGGAACAACAACAACCAATTCCTCAAGATGCTGTGACT 6900
Qy 961 GAGCACAATGGAACCGACCCCAACAACAACAAGCAAGCTCTCTCAACAATACTAATAACA 1020
Db 6901 GAACTTTGACAAAAATAACAACATGCAACACGCTCCATGCGCCCTCATATAACATACCAACA 6960
Qy 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCACCGCGAAC 1080
Db 6961 ATCTCTACTAACAACACTCTCAACAACACTTCAGCACTCTCTCTGACCACTTACAAAC 7020
Qy 1081 ATCAACCAATAATGATACACAACAGTGAACTAGCAGAAAGCGAACCAACCAATGCTCAGTTG 1140
Db 7021 ACCACCAATGACAACAACAGCAGCAATCACTGAATATGAGCAAAACCCAGTGCCCTCG 7080
Qy 1141 AACCAACTCTAGATCCAAAGAAAATCCACACAGCAGCAAGACCAACAAGCAGCAAC 1200
Db 7081 ATAACAACCTGCTCCCAACGCGAAATCCACACAGCAAGACCAACAGCAGCAAAAAA 7140
Qy 1201 AACATCATCATGACATCATGATATACAGCAACCAACCAACCAATTTCTCTCCGAT 1260
Db 7141 GGCCCGCCACAACGGCACCAACACGAAATGAGCAATTTACAGCTCTCCCGCCAC 7200
Qy 1261 TCTAGTCGCAACCGCGCTCTCTATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
Db 7201 CCCAGCTGACGTCACACACTCTGTATATTTAGAGAAAGCGAAGTATCTCTGGAGG 7260
Qy 1321 GAAGCGCATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAAATGATTTTGACCA 1380
Db 7261 GAAGCGCATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAAATGATTTTGACCA 7320
Qy 1381 GTTCCAAATACAAAACAACTTTGATGAATCTCTAGTTCTGGTGGCTCGGCTGAGAA 1440
Db 7321 GTTCCAAATACAAAACAACTTTGATGAATCTCTAGTTCTGGTGGCTCGGCTGAGAA 7380
Qy 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTTCTTAAATATAAATGAG 1500
Db 7381 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTTCTTAAATATAAATGAG 7440
Qy 1501 AACACTGCTACTCTGGAGAAAATGAGAAATGATGTGATGAGAGTTAAGAAATTTGGAGC 1560
Db 7441 AACACTGCTACTCTGGAGAAAATGAGAAATGATGTGATGAGAGTTAAGAAATTTGGAGC 7500
Qy 1561 GTTCAAGGAGATGACTCGCGCAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAT 1620
Db 7501 GTTCAAGGAGATGACTCGCGCAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAT 7560
Qy 1621 GAAGGACTTTACACTGCTGTTTAAATTAATAATCAAAACAATTTGGTCTGCAGGTTGAGG 1680
Db 7561 GAAGGACTTTACACTGCTGTTTAAATTAATAATCAAAACAATTTGGTCTGCAGGTTGAGG 7620
Qy 1681 CGTCTAGCAATCAAACTGCCAAATCTTTGGAACTCTTTATTTGAGAGTCACAACTGAGAA 1740
Db 7621 CGTCTAGCAATCAAACTGCCAAATCTTTGGAACTCTTTATTTGAGAGTCACAACTGAGAA 7680
Qy 1741 AGAACATTTCTCTTAATCAATAGACATGCTATTTGATCTTTACTCAAGAATGGGAGGA 1800
Db 7681 AGAACATTTCTCTTAATCAATAGACATGCTATTTGATCTTTACTCAAGAATGGGAGGA 7740
Qy 1801 ACATCAAGTCTTTGGACTGATTTGATCGGGATAGAGACTCTGTCACAAATATTT 1860
Db 7741 ACATCAAGTCTTTGGACTGATTTGATCGGGATAGAGACTCTGTCACAAATATTT 7800
Qy 1861 TCAGAGCAAAATGACCAAAATTAAGGACGCAACAAAAGAGGGGACTGGTTGGGGCTG 1920
Db 7801 TCAGAGCAAAATGACCAAAATTAAGGACGCAACAAAAGAGGGGACTGGTTGGGGCTG 7860
Qy 1921 GGTGTAAATGTTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTTGCTACTA 1980

Db 7861 GGTGGTAATGTTGGACATCCGACTGGGTGTTCTTACTAACTTGGCAATTTTGCTACTA 7920
Qy 1981 TTATCCATAGCTGTCTTGATTTGCTCTATCTGCTATCTGCTATCTTTTACTTAATAATATC 2040
Db 7921 TTATCCATAGCTGTCTTGATTTGCTCTATCTGCTATCTGCTATCTTTTACTTAATAATC 7980
Qy 2041 GGATAA 2046
Db 7981 GGATAA 7986
RESULT 8
AX717721
LOCUS AX717721 2051 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 9 from Patent WO02079239.
ACCESSION AX717721
VERSION AX717721.1 GI:29890729
KEYWORDS
SOURCE Marburg virus (strain Musoke)
ORGANISM Marburg virus (strain Musoke)
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1
AUTHORS Grogan, C. C., Hevey, M. C. and Schmaljohn, A. L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 9 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source 1. .2051
/organism="Marburg virus (strain Musoke)"
/mol_type="unassigned DNA"
/db_xref="taxon:33727"
/note="chimeric molecule between Marburg virus
Glycoprotein 1 and Marburg virus Glycoprotein 2"
ORIGIN
Query Match 72.6%; Score 1485.6; DB 6; Length 2051;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1705; Conservative 0; Mismatches 329; Indels 12; Gaps 1;
Qy 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAAATCTCCCT 60
Db 10 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTAATTCAGGACCAAAAATCTCCCC 69
Qy 61 GTTTTAGAAATGTCTAGTAACAGCAACCTCAAGATGTAGATTGAGTGTCTCCGGAAC 120
Db 70 ATTTTAGATAGTAGTAGTAATAATCAACCCCAAAATGTGGATTCTGCTATGCTCCGGAAC 129
Qy 121 CTCGAAAGACAGAAGATGTTTCATCTGATGGGATTTACATGAGTGGCAAAAAGTTGCT 180
Db 130 CTCGAAAGACAGAAGACGCTCCATCTGATGGGATTCACATGAGTGGCAAAAAGTTGCT 189
Qy 181 GATTCCCTTTGGAAGCATCTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 190 GATTCCCTTTGGAAGCATCTCAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 249
Qy 241 GTTCAGTATACGGAAGGAGAAAGCCAAACATGTTTCAATATAAGTGTAAACGACCT 300
Db 250 GTTCAGTATACGGAAGGAGGAGGAGCAACCAACATGCTACATATAAGTGTAAACGATCC 309
Qy 301 TCTCGAAAATCTCTGCTGCTGATCTCCCGAGTAATATCCGCGATTACCTTAATATGTA 360
Db 310 TCTCGAAAATCTCTGCTGCTGATCTCTTACCAACATCCGCTGACTATCCGAAAATGCAA 369
Qy 361 ACTCTTCATCATATTCAGGTCAAAACCTCATCATCAGGAGGATTCCTCCATTTG 420
Db 370 ACTATCCATCATATTCAGGTCAAAACCTCATCATCAGGAGGATTCCTCCATTTATG 429
Qy 421 GGGGCAATTTTCTTGATGATCGGTTGCTCTCAACAATGTTTACCGAGGCAAGTCTTC 480
Db 430 GGAGCAATTTTCTTGATGATCGCAAT-----ATGTACCGAGGCAAGTCTTC 477
Qy 481 ACTGAAGGAAATATAGCAGCTATGATTGTTAATAAGACAGTTTCACAGAAATGATTTTCT 540

||||| 478 ACTGAAGGAAACATAGCAGCTATGATTGTCAATAAGACAGTGCCAAAAATGATTTTCG 537
||||| 541 AGGCAAGGACAGGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTGGACAGC 600
||||| 538 CGGCAAGGACAGGTTACCGTCATATGATCTGACTTCTACTATATATATTTGGACAGT 597
||||| 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTGGGATCTTCCAAAGAAATACAACTCC 660
||||| 598 AGTACCGAAGCGCAACGAATGACACTGGATGTTTCCGGCGCTTTTCAAGAATACAAATTCT 657
||||| 661 ACAACAACTCAAAATGCTCCATCTCTTAACCTCCATCCCTGCCACACGTAACCTCG 720
||||| 658 ACAAGAAACCAACATGTGCTCCGTCCTCAAAATACCTCCACCACTGCCACAGCCGCTCG 717
||||| 721 AGCATCTCACTTACAAATTAATCTCAAAATTAATCTGTAATCTTGAACCTATGAAACCCAAGT 780
||||| 718 GAGATCAAACTCACAAAGCAACCCCACTGATGCCCAAACTCAATACCAAGCAACCAAGC 777
||||| 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAGAACAGGGGCCCCACACA 840
||||| 778 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAAACCCACACA 837
||||| 841 ACTCTTAATGTAGTCACTGAAACAGAAACAAATCGTCAACAAATTTGTCCACTCCTTCACTA 900
||||| 838 ACTTCTGATGGGTCACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACCA 897
||||| 901 CATCCAAGCAGCTCACAAATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAACT 960
||||| 898 CAACCAAGCAGCCACAGGAGGAGAAACAAACAACCACTTCCCAAGATGCTGTGACT 957
||||| 961 GAGCAATGGAACCGACCCCAACAACAACAGCAACCGCTCTCAACAATACTAATACA 1020
||||| 958 GAACTAGACAAAATAACACACTGACAAACCGTCCATGCCCTCTATAACACTACCACA 1017
||||| 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCGGCAAC 1080
||||| 1018 ATCTCTACTAAACACCTCCAAACACAACTTCAGCACTCTCTTGCAACCAATTAACAAAC 1077
||||| 1081 ATCACCAATATGATACAAACGTAAGTACGAGAAAGCAACCAACCAATGCTCAGTTG 1140
||||| 1078 ACCACCAATGACAAACACAGAGCAATCACTGAAATGAGCAACCAAGTGCCTCG 1137
||||| 1141 AACACAATCTTAGATCCAAAGAAAATCCCAACAGAGCAAGACCAACAGCAACAC 1200
||||| 1138 ATAAACAACCTGCTCCAAAGGAAATCCCAACAGCAAGAGACCAAGAGCAAGCAAAAA 1197
||||| 1201 AACATCATGACGACATCAGATATACAGCAACACCCCAACAAATCTTCTCCGAT 1260
||||| 1198 GGGCCCGCCCAACCGGCAACCAACGCAATGAGCATTTTCCAGCTCCTCCCGCCACC 1257
||||| 1261 TCTAGTCCGACAAACCGGCTCTTATATATCTTAGAAAGAACGATCGATCTCTGGAGG 1320
||||| 1258 CCAGCTGAGTGCACAACTCTTGATATTTTCAAGAGAGGAGATCGATCTCTGGAGG 1317
||||| 1321 GAAGCGCATGTTCCCTTTCTGGATGGGTTAAATAATGCTCCCAATGATTTGACCCA 1380
||||| 1318 GAAGCGCATGTTCCCTTTCTGGATGGGTTAAATAATGCTCCCAATGATTTGACCCA 1377
||||| 1381 GTTCCAAATACAAAACAACTTTTGAATGATCTAGTTCTGGTGCCTCGCTGAGAA 1440
||||| 1378 GTTCCAAATACAAAACAACTTTTGAATGATCTAGTTCTGGTGCCTCGCTGAGAA 1437
||||| 1441 GATCAACATGCTCCCAATATAGTTTAACTTATCTTATTTTCCCTAAATAAATGAG 1500
||||| 1438 GATCAACATGCTCCCAATATAGTTTAACTTATCTTATTTTCCCTAAATAAATGAG 1497
||||| 1501 AACATGCTTACTCTGGAGAAAATCAGAAATGATGTGATGACAGTTAAGAAATTTGGAGC 1560
||||| 1498 AACATGCTTACTCTGGAGAAAATCAGAAATGATGTGATGACAGTTAAGAAATTTGGAGC 1557
||||| 1561 GTTCAGGAGGATGACTGGCCGACGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAATT 1620
|||||

Db 1558 GTTCAGGAGGATGACCTGGCCGACAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAATT 1617
Qy 1621 GAAGGACCTTACACTGCTGTTTTTAATTAATAAATAAATAAATAAATAAATAAATAAATAA 1680
Db 1618 GAAGGACCTTACACTGCTGTTTTTAATTAATAAATAAATAAATAAATAAATAAATAAATAA 1677
Qy 1581 CGTCTAGCCCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 1740
Db 1678 CGTCTAGCCCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 1737
Qy 1741 AGAACATTTCTCTTAATCAATAGACATGCTATTGCACTTTCTACTCAAAAGATGGGAGGA 1800
Db 1738 AGAACATTTCTCTTAATCAATAGACATGCTATTGCACTTTCTACTCAAAAGATGGGAGGA 1797
Qy 1801 ACATGCAAAAGTCTTGGACCTGATTGTTGCAATCGGATAGAACTTGTTCAAAAAATATT 1860
Db 1798 ACATGCAAAAGTCTTGGACCTGATTGTTGCAATCGGATAGAACTTGTTCAAAAAATATT 1857
Qy 1861 TCAGAGCAAAATTTGACCAATTTAAAAGGACGAAACAAAAGAGGGGACTGTTGGGCTCTG 1920
Db 1858 TCAGAGCAAAATTTGACCAATTTAAAAGGACGAAACAAAAGAGGGGACTGTTGGGCTCTG 1917
Qy 1921 GGTGTTAAATGTTGGACATCCGACTGGGCTGTTCTTACTAACTTTGGGCAATTTTGCTACTA 1980
Db 1918 GGTGTTAAATGTTGGACATCCGACTGGGCTGTTCTTACTAACTTTGGGCAATTTTGCTACTA 1977
Qy 1981 TTATCCATAGCTGCTTTGATTGCTCTATCTCTGATATTTGTCGATCTTTTACATAATATATC 2040
Db 1978 TTATCCATAGCTGCTTTGATTGCTCTATCTCTGATATTTGTCGATCTTTTACTAAATATATC 2037
Qy 2041 GGATAA 2046
Db 2038 GGATAA 2043

RESULT 9
AF005735
LOCUS
DEFINITION Marburg virus strain M/Germany/Marburg/1967/Ratayczak glycoprotein precursor (GP) gene, complete cds.
ACCESSION AF005735
VERSION AF005735.1 GI:2459879
KEYWORDS
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales; Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 1 to 2046)
Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and Feldmann,H.
Variation in the glycoprotein and VP35 genes of Marburg virus strains
JOURNAL Virology 240 (1), 138-146 (1998)
MEDLINE 98110148
PubMed 9448638
REFERENCE 2 (bases 1 to 2046)
Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
Direct Submission
Submitted (29-MAY-1997) Special Pathogens Branch, Centers for Disease Control and Prevention, 1600 Clifton Road, Building 15, Room SB611, Mail Stop G14, Atlanta, GA 30333, USA
FEATURES
Location/Qualifiers
1. .2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/Germany/Marburg/1967/Ratayczak"
/db_xref="taxon:11269"
1. .2046
/gene="Gp"
1. .2046
/gene="Gp"
/note="virion surface glycoprotein"
/codon_start=1

```

/product="glycoprotein precursor"
/protein_id="AAC40460.1"
/db_xref="GI:2459880"
/translation="MKTTCLFLSILILIQIKTLPILEIASNNQPNVDSVCSGLQKT
EDVLMGPTLSQVADSPLEBSKRWAFRTGVPKPNVEYTEGEAEKTCYNI SVTDPBG
KSLLDLPNTIRDYPKCKTHI IQQNPFAQIALHLWGAPFLYDRIASITVMYGRVF
TEGNIAMVUNKTVHMKI FSRQGGYRHMILTSNKYWTNSNGTQTNDCGFCGALQEV
NSTKNQCAPSKI SPPLTPARPEIKPTSTPDATILNTDNDNDEBLI ISGSSGSGQ
EPTYSOAVTKQGLSSMPPTSPSPQSPQEGNNTHDSQGTVEPEKNTINTTAQSPMP
PHNTTAISTNNSTKNSFSLVSQNTNTNDYDQSTATENBOQTSAPSKTTLPPPTGNLT
AKSNNTKPTPTTAPNTNTHDSSSPPTNPPTQHLVYFRKRSILWRGDMFPLDGL
LINAPIDPVPNTKTI FDESSSGSAEEDOHASPNISLTSYFPNINENTAYSGEN
ENDDAELRIWSVOEDDLAAGLSMI PFPGPGIEGLYTAGLIKNNNLVCLRLRANOT
AKSLELLRLVTEBTEPSLINRHAIDPELLFRWGCTCKVLGPDCCIGIEDLSRNISEQI
DOI KDRQKEGTGWLGGKMWISDWGVTNLNLIILLISIAVLIALSICRIPTYIIG"
1..81
/ gene="Gp"
82..2043
/ gene="Gp"
/product="glycoprotein"
1930..2019
/ gene="Gp"
/ note="encodes transmembrane anchor peptide"

```

ORIGIN

```

Query Match 72.1%; Score 1474.8; DB 14; Length 2046;
Best Local Similarity 82.6%; Pred. No. 0;
Matches 1689; Conservative 0; Mismatches 357; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTCTCATGTAGTCTCTATTTTAACTCCAAAGTATATAAACTCTCCCT 60
DB 1 ATGAAGACCATATGTCCTTTATCATGTCTTATCTTTATTAATCCAAAGGATATAAACTCTCCCT 60

QY 61 GTTTTAGAAATGTCTAGTAACAGCCAACTCAAGATGTAGATTTCAGTGTGCTCCGGAAAC 120
DB 61 ATTTTAGAGATAGCTAGTAACAACTCAACCCCAAAATGTGGATTTCGGTATGCTCCGGAACT 120

QY 121 CTCGAAAGACAGAAAGATGTTTCATCTGATGGGATTTACATCTGAGTGGGCAAAAATGTGCT 180
DB 121 CTCGAAAGACAGAAAGATGTTTCATCTGATGGGATTTACATCTGAGTGGGCAAAAATGTGCT 180

QY 181 GATTCCTCTTTGGAGCATCTTAAGCATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
DB 181 GATTCCTCTTTGGAGCATCTTAAGCATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240

QY 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACAATATAAGTGTAACAGACCCCT 300
DB 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACAATATAAGTGTAACAGGATCCC 300

QY 301 TCTGGAATACTCTGCTCTGGATCCTCCAGTAAATATCCGCGATTACCTTAATGTAAA 360
DB 301 TCTGGAATACTCTGCTCTGGATCCTCCAGTAAATATCCGCGATTACCTTAATGTAAA 360

QY 361 ACTGTTTCATCATATTCAGGTCAAAACCTCATGCAAGGGGATGCGCTCATTTGTGG 420
DB 361 ACTATCCATCATATTCAGGTCAAAACCTCATGCAAGGGGATGCGCTCATTTGTGG 420

QY 421 GGGCAATTTTCTGTATGATCGCTCTCAACAATGTACCGAGGCAAGTCTTC 480
DB 421 GGAGCAATTTTCTGTATGATCGCTCTCAACAATGTACCGAGGCAAGTCTTC 480

QY 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAGACAGTTTCACAGAAATGATTTTTCT 540
DB 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAGACAGTTTCACAGAAATGATTTTTCT 540

QY 541 AGGCAAGGCAAGGTATTCGTCACATGAATTCGACCTCCACCAATATAATATTTGACAGC 600
DB 541 AGGCAAGGCAAGGTATTCGTCACATGAATTCGACCTCTACTAATAATAATATTTGCAAGT 600

QY 601 AGCAATGAAACGACAGAGAAATGATACGGGATGTTTGGCATCCTCCAAAGATATCAACTCC 660
DB 601 AACAATGGACAAACGAAATGACACTGGATGCTTTCGGTGCTTTCAGAAATATCAACTCC 660

```

```

QY 661 ACAACAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACCTCG 720
DB 661 ACAAGAATCAAAATGCTGCTCCGTCCAAATACCTCCACACTGCGCCACAGCCCGTCCA 720

QY 721 AGCATTCACCTCTCAAAATCTCAAAATTAATCTGCTAAATCTGGAACATATGAACCCCAAGT 780
DB 721 GAGATCAAAACCCACAGCACCCCACTGATGCGCACCACTCAACACCAAGACCCCAAC 780

QY 781 AGGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
DB 781 AATGATGATGAGGACCTCATAAATCATCCGTTCCAGGGTCCGAGAACAGGAAACCCCTATACA 840

QY 841 ACTCTTAATCTGCTGCTGACGACGAAACAACTCGTCAACAATATTTGTCACCTCTCTCACTA 900
DB 841 ACTTCAGATGCGGTCTCTTAAGCAGGCTTTTCATCAACAATGCGCCACCTCCCTCACCA 900

QY 901 CATCCAGCACCTCACAACAATGAGCAAAAACAGTACGAATCTCTCCGACATGCTGTAACT 960
DB 901 CAACCAAGCAGCCACAGCAAGAAAGAAACACACACAGACCAATTCCTCAAGGTACTGTGACT 960

QY 961 GAGCACAATGGAACCGACCCCAACAACAACAGCAACCGTCTCTCAACAATACTAATAACA 1020
DB 961 GAACCCAAACAAACCAACAACAACCTGCAACCGTCCATGCGCCGCCCAACAACCACTGCA 1020

QY 1021 ACTCCGACCTATTAACACTCTCAAGTACAACTCTGAGTACTCTCTCCCTCCGACCGCAAC 1080
DB 1021 ATCTCTACTAACAACACTCTCCAGAACAACTCTGAGCACTCTCTCTGATCTCACTACAAAC 1080

QY 1081 ATCAACCAATATGATACACAACAGTGAACCTAGCAGAAAGCGAAACAAACCAATGCTCAGTTG 1140
DB 1081 ACCACCAATTTAGCACACAGAGCAGCAGCAGCACTGAAATGNAACAAACAGTGGCCCTCG 1140

QY 1141 AACACAATCTTAGATCCAAACAGAAAATCCCAACACAGGACAAAGACACCAACAGACCAACC 1200
DB 1141 AAAACAACCTCTCCCTCCCAACAGGAAATCTTACACAGCAAAAGAGACAAACAACAGGAA 1200

QY 1201 AACATCATCATGACGACATCAGATATAAACAAGCAAAACACCCCAACAATTTCTTCCCGAT 1260
DB 1201 GGGCCCAACCAACGGGCAACCAATATGACAAATGGGCAATTTAACAGTCTCTCCGCCACC 1260

QY 1261 TCTAGTCCGACAAACCGCTCTCTATATATCTTTAGAAAGAAACCATGATCGATCTCTCGAGG 1320
DB 1261 CCAACCCGACCAACAACATCTTTGATATATTTAGAAAGAAACAGAGTATCTCTCGAGG 1320

QY 1321 GAAGGCGACATGTTTCTCTGATGGGTAAATAAATGCTCCAATTTGATTTTGACCCA 1380
DB 1321 GAAGGCGACATGTTTCTCTGATGGGTAAATAAATGCTCCAATTTGATTTTGATCCA 1380

QY 1381 GTTCCAAATACAAAACAAATCTTTGATGAAATCTCTAGTCTGCTGCTCGGCTGAGGAA 1440
DB 1381 GTTCCAAATACAAAACAAATCTTTGATGAAATCTCTAGTCTGCTGCTCGGCTGAGGAA 1440

QY 1441 GATCAACATGCTCCCAATATTTAGTTTAACTTTTATCTTATTTCTTAATAATAAGTAA 1500
DB 1441 GATCAACATGCTCCCAATATTTAGTTTAACTTTTATCTTATTTCTTAATAATAAGTAA 1500

QY 1501 AACACTGCTACTCTCGAGAAATAGAGATGATTTGATGACAGAGTTAAGAAATTTGGAGC 1560
DB 1501 AACACTGCTACTCTCGAGAAATAGAGACGATTTGATGACAGAGTTAAGAAATTTGGAGC 1560

QY 1561 GTTCAGAGGATGACCTGGCGCAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 1620
DB 1561 GTTCAGAGGATGACCTGGCGCAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 1620

QY 1621 GAAGGACTTTACACTGCTGTTTTTAATAAATCAAAACAAATTTGGTCTGCGAGTTCAGG 1680
DB 1621 GAAGGACTTTATCTGCTGTTTTTAATAAATCAAAACAAATTTGGTCTGCGAGTTCAGG 1680

QY 1681 CGTCTAGCCAACTCAAACTGCGCAATCTCTTATTTGAGAGTCACACTGAGGAA 1740
DB 1681 CGTCTAGCCAACTCAAACTGCGCAATCTCTTATTTGAGAGTCACACTGAGGAA 1740

QY 1741 AGAACATTCCTTTAATCAATAGACATGCTATTGACTTTTCTACTCAACAAGATGGGGAGGA 1800

```

Db	1741	AGGACATTTTCCTTAATTAATAGACATGCCATTGACCTTTCTACTCAAAAGGTGGGGAGGA	1800			
Qy	1801	ACATGCAAAAGTCTGTGGACCTGATTGTTCATCGGGATAGAAGACTTGTCCAAAAATATT	1860			
Db	1801	ACATGCAAAAGTCTGTGGACCTGATTGTTCATCGGGATAGAAGACTTGTCCAGGAATATT	1860			
Qy	1861	TCAGAGCAAAATTGACCAAAATTAATAAAGGACGAAACAAAGAGGGGACTGGTTGGGGTCTG	1920			
Db	1861	TCGGAAACAAATTGACCAAAATCAAAAAAGATGAACAAAAAGAGGGGACTGGTTGGGGTCTA	1920			
Qy	1921	GGTGGTAATAGGTGGACATCCGACTGGGGTGTCTTACTTAACCTTGGGCATTTTGTCTACTA	1980			
Db	1921	GGTGGTAATAGGTGGACATCCGACTGGGGTGTCTTACTTAACCTTGGGCATTTTGTCTACTA	1980			
Qy	1981	TTATCCATAGCTGCTTGTGATTGCTCTATCCCTGATTTTGTCTATCTTTACTAAATATATC	2040			
Db	1981	TTATCCATAGCTGCTTGTGATTGCTCTATCCCTGATTTTGTCTATCTTTTACCAATATATC	2040			
Qy	2041	GGATAA	2046			
Db	2041	GGGTAA	2046			
RESULT 10						
MAVSPAA						
LOCUS						
DEFINITION	Marburg Virus genomic RNA of GP gene.					
ACCESSION	X68493					
VERSION	X68493.1 GI:296960					
KEYWORDS	gp gene; structural protein.					
SOURCE	Lake Victoria marburgvirus					
ORGANISM	Lake Victoria marburgvirus					
Viruses; ssRNA negative-strand viruses; Mononegavirales;						
Filoviridae; Marburg-like viruses.						
REFERENCE	1 (bases 1 to 2948)					
AUTHORS	Bukreyev.A., Volchkov.V.E., Blinov.V.M. and Netesov.S.V.					
TITLE	The GP-protein of Marburg virus contains the region similar to the 'immunosuppressive domain' of oncogenic retrovirus P15E proteins					
JOURNAL	PRES Lett. 323 (1-2), 183-187 (1993)					
MEDLINE	93265932					
PUBMED	8495737					
REFERENCE	2 (bases 1 to 2948)					
AUTHORS	Bukreyev.A.A., Volchkov.V.E., Blinov.V.M., Dryga.S.A. and Netesov.S.V.					
TITLE	The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain					
JOURNAL	Arch. Virol. 140 (9), 1589-1600 (1995)					
MEDLINE	96028047					
PUBMED	7487490					
REFERENCE	3 (bases 1 to 2948)					
AUTHORS	Bukreyev.A.A.					
TITLE	Direct Submission					
JOURNAL	Submitted (28-SEP-1992) A.A. Bukreyev, All-Union Inst. of Molecular Biology, NPO 'Vector', Koltsovo, Novosibirsk region, 633159, USSR					
FEATURES	Location/Qualifiers					
source	1..2948					
	/organism="Lake Victoria marburgvirus"					
	/mol_type="genomic RNA"					
	/strain="Popp"					
	/db_xref="taxon:11269"					
	6..2851					
gene	/gene="Gp"					
mRNA	(6.17)..(2841.2851)					
	/gene="Gp"					
	122..2167					
CDS	/gene="Gp"					
	/codon_start=1					
	/product="structural protein"					
	/protein_id="CAA48507.1"					
	/db_xref="GI:296961"					
	/db_xref="GOA:P35254"					
	/db_xref="UniProt/Swiss-Prot:P35254"					

/translation="MKTTCLFISLILIOGKITLPILIEIASNNQPNQVDSVCSGTLQK EDVHLMGFTLSQKVADEASKEAWFAFTPPKKNVEYTEGEAEACTCNIISVTDPS KSLILLDPKPTNIIRHIQONPHAQGIALHMGAFFLYDIRASTMYRVRGLOEY TEGNIAMTVNKTVHKMI FSRQOGYRHNMTSTNKVYTTNNQNTDNTGFCFALOEY NSTKNQTCAPLSPSPLPATRPBKPTSTPTDATTLLTDPNNDDDELITCSGSGSGEO EPTYSDDATVKQGLSMTMPTSPBPSTPQQEGNWDHSQGTVEPNKNTYTAQSPMP PHNTAISTNNSTKNPNTSLVSLQNTNYDTQSTATENEQSDPSKTLTPPGNLT AKSTNNQGTPTTANMTNGHILTSPTNPNTTQHLVYPRKRSILMRGDMPPFLDG LNPADIDFQVNTKNTIIPDESSSGASAEEDQHASPNISLITSLYPNNVENTRAYSGEN ENDCDAELRISWVOEDDLAAGLSWI PFPGGIEGLYTAGLIGNONNKLRRRLANOT AKSLELLLRVTTEERTPSLINRHADIFLPLTRMGGTCVKLGPDCCTGIEDLSRNISEQI DQIKODEQEGTGWGLGRKWTSDMGLVTLNLGILLLSIAVLIALSICRIFTKYIG"									
ORIGIN		Query Match	72.0%;	Score 1473.2;	DB 14;	Length 2948;			
		Best Local Similarity	82.5%;	Pred. No. 0;					
		Matches 1688;	Conservative	0;	Mismatches	358;	Indels	0;	Gaps
Qy	1	ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT	60						
Db	122	ATGAAGACCACATGTCTCTTTATCAGTCTTATCTTAATCCAAAGGATATAAAACTCTCCCT	181						
Qy	61	GTTTTAGAATTTCTGTAGTAACAGCCAACTCAGATGTAGATTCAGTGTGCTCCGGAAC	120						
Db	182	ATTTTAGAGATAGCTAGTAAACAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT	241						
Qy	121	CTCAAAAGACAGAAGATGTTTCACTGTGGGATTTTACACTGAGTGGGCAAAAGTTGCT	180						
Db	242	CTCAGAGNACAGAAGATGTCATCTGTGGGATTCACACTGAGTGGGCAAAAGTTGCT	301						
Qy	181	GATTCCCTTTGGAAGCATCTAAACGATGGGTTTCAGGACAGGTGTTCTCTCCAGAAG	240						
Db	302	GATTCCCTTTGAGGCATCCAAAGCATGGGCTTCAGGACAGGTGTACCTCCCAAGAAT	361						
Qy	241	GTTGAGTATACGGAAGAGAGAGCAAAACATGTTACATATAGTGTAAACAGACCCT	300						
Db	362	GTTGAGTATACAGAGGGGAGGAAGCAAAACATGCTCAATATAGTGTAAACGGATCCC	421						
Qy	301	TCGTGGAATATCCCTTGCTGTGGATCTCCAGTAAATACCGCATTAACCGATTAACCT	360						
Db	422	TCGTGGAATATCCCTTGCTGTGGATCTCTCTA CCAACATCCGTGACTATCTCTAAT	481						
Qy	361	ACTGTTTCATCATATTTCAAGGTCAAAACCCCTCATGCA CAGGGATTCGCCCTCCA	420						
Db	482	ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCGCAAGGATCGCCCTCCATT	541						
Qy	421	GGGGCATTTTTCTGTGATGATCGCGTTGGCTCTACAAACATGTACCGAGCAAGGTCTTC	480						
Db	542	GGAGCATTTTTCTGTGATGATCGCATTTGCCCTCCACAAACATGTACCGAGGACAGTCTTC	601						
Qy	481	ACTGAAGGAAATATAGCAGCTATGATTTGTTTAAATAGACAGTTTCACAGAAATGATTTT	540						
Db	602	ACTGAAGGAAACATAGCAGCTATGATTTGTTCAATAGACAGTGCACAAATATGATTT	661						
Qy	541	AGGCAAGGACCAAGTGTATCGTCA CATTGAACCTTGACCTCCA CCAATAATATTGGCA	600						
Db	662	AGGCAAGGACAGGGTACCGTCA CATTGAACCTTGACCTCCA CCAATAATATTGGCA	721						
Qy	601	AGCAATGAACGAGAGAAATGATACGGGATGTTTTGGGATTCCTCCAAGAAATACAATCC	660						
Db	722	AACAATGGAACACAAACGAATGACACTGGATGCTTCGGTGCTCTTCAAGAAATACAATCC	781						
Qy	661	ACAAACANTCAACATGCCCTCCATCTTTAAACCTCCATCCCTGCCCAACAGTAACTCCG	720						
Db	782	ACGAAGAATCAAAACATGTGCTCGGTCCAAATACCCTCACCTGCCCCACAGCCCCGTCCA	841						
Qy	721	AGCATTCACCTCTACAAATACTCAAATTAATPACTGTCTAAATCTGGAACTATGAACCCCA	780						
Db	842	GAGATCAAAACCCACAGCACCCCAACTGATGCCACCACTCAACACCCACAGACCCCAAC	901						
Qy	781	AGCGAGATGAGACCTTATGATTTCCGGCTCAGGATCTGGAGAAACAGGGGCCCCACACA	840						
Db	902	AATGATGATGAGACCTCTAAACATCCGGTTCAGGGTCCGAGAAACAGGAACCCCTATACA	961						

```
Qy 841 ACTCTTAATGTAGTCACTGAACAGAAACATCGTCAACAATATTTGTCTCACTCTCTACTA 900
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 962 ACTTCAGATGGGTCACTTAAGCAAGGGCTTTTCATCAACAATGCGCCACCTCCCTCA 1021
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 901 CATCCAGCACTCTCAACAATGAGCAAAACAGTAGTCAATCTCTCCGACATGCTGTA 960
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1022 CAACCAAGCAGCCACAGCAAGAGGAAACACACAGACCATTCCTCAAGAGTACTGT 1081
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 961 GAGCAATAGGAACCGACCCCAACACACACAGCAACAGTCTCTCAACAATACTAATA 1020
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1082 GAACCCACAAACCAACACACAGCGCAACCGTCCATGCGCCGCCCAACACACACTGCA 1141
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1021 ACTCCCACTTAACACTCTCAAGTACCACTCTAGTACTCTTCTCCCTCCACCGCAAC 1080
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1142 ATCTCTACTAACAACCTCCCAAGCAACACTTTCAGCACCTCTCTGTATCACTACA 1201
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1081 ATCAACCAATATGATACACAACGTGAACCTAGCAGAAAGCGAAACCAACCAATGCT 1140
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1202 ACACCAATAGCAGACACACAGAGCACAGCCACTGAAATGAAACCAACGAGTCC 1261
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1141 AACCAACTCTAGATCCCAAGAAATCCCAACGAGACAGACACCAACAGACCAAC 1200
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1262 AAAACAACCTCGCTCCCAACAGGAAATCTTACCACAGCAAGAGCACTTAACAAC 1321
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1201 AACATCATGACGACATCAGATATACAAAGCAACACCCCAACAATTTCTTCCGG 1260
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1322 GGGCCCAACCAACGGCACCAAAATATGACAAATGGCAATTTAAACGAGTCCCTCC 1381
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1261 TCTAGTCGACAAACCGCCCTCTTATATCTTATAGAAAGAACGATCGATCTCTGG 1320
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1382 CCCAACCGGACCAACAACATCTTGTATATTTGAAAGAAACGAATATCTCTGG 1441
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1321 GAAGCGACATGTTCCCTTTCTCGATGGGTATATAATGCTCCAATGATTTTGAC 1380
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1442 GAAGCGACATGTTTCTTTCTCGACGGGTATATAATGCTCCAATGATTTTGAT 1501
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1381 GTTCCAAATACAAAACATCTTTGATGAATCTCTAGTTCGTGCTCGGCTGAGG 1440
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1502 GTTCCAAATACAAAACATCTTTGATGAATCTCTAGTTCGTGCTCGGCTGAGG 1561
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTCTTAATATAAATG 1500
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1562 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTCTTAATATAAATG 1621
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1501 AACACTGCTACTCTGGAGAAATGAGAAATGATGTGATGAGAGTTAAGAAATTTG 1560
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1622 AACACTGCTACTCTGGAGAAATGAGAAATGATGTGATGAGAGTTAAGAAATTTG 1681
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1561 GTTCAGGAGGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTTTTGGCCCTG 1620
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1682 GTTCAGGAGGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTTTTGGCCCTG 1741
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1621 GAAGGACTTTACATGCTGCTTTTAAATTAATAATCAAAACATTTGCTCAGGTT 1680
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1742 GAAGGACTTTATCTGCTGGTTTTAATTAATAACCAAAACAATTTGGTCTCAG 1801
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1681 CGTCTAGCAATCAAACTGCCAAATCTCTTGGAACTCTTATTTGAGAGTCACAACT 1740
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1802 CGTCTAGCAATCAAACTGCCAAATCTCTTGGAACTCTTATTTGAGAGTCACAA 1861
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1741 AGAACAATTTCTTAATCAATAGACATGCTATTTGACTTTCTCTACAAGATGGG 1800
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1862 AGGACATTTTCTTAATTAATAGACATGCCATTTGACTTTCTCTACAAGATGGG 1921
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1801 ACATGCAAGTCTTGGACCTGATTTGATCGGGATAGAGACTGTCTCAAAAATATT 1860
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1922 ACATGCAAGTCTTGGACCTGATTTGATCGGGATAGAGACTGTCTCAAAAATATT 1981
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1861 TCAGAGCAATTTGACCAATTTAAAGAGCAACAAAGAGGAGGACTGTGTGGGCTG 1920
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1982 TCGNACAAATTTGACCAATTTAAAGAGCAACAAAGAGGAGGACTGTGTGGGCTG 2041
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Qy 1921 GGTGTAATGTGGATCGGATCCGACTGGGTGTTCTTACTAACTTGGCACTTTTCTACTA 1980
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2042 GGTGTAATGTGGATCCGACTGGGTGTTCTTACTAACTTGGCACTTTTCTACTA 2101
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1981 TTATCCATAGTGTCTTGATTGCTCTATCTGCTATTTGTGCTTCTTAAATATATC 2040
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2102 TTATCCATAGTGTCTTGATTGCTCTATCTGCTATTTGTGCTTCTTAAATATATC 2161
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2041 GGATAA 2046
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2162 GGGTAA 2167
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
MVIRPR
LOCUS 19112 bp RNA linear VRL 22-JAN-1996
DEFINITION Marburg virus (Popp) NP, VP35, VP40, GP, VP30, VP24, L genes.
ACCESSION Z29337
VERSION Z29337.1 GI:450908
KEYWORDS GP protein; L protein; NP protein; VP24 protein; VP30 protein; VP35 protein; VP40 protein.
SOURCE Lake Victoria marburgvirus
ORGANISM Lake Victoria marburgvirus
VIRUSES: ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE 1 (bases 2851 to 5818)
AUTHORS Bukreyev,A.A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.
TITLE The VP35 and VP40 proteins of filoviruses. Homology between Marburg and Ebola viruses
JOURNAL PNAS Lett. 322 (1), 41-46 (1993)
MEDLINE 93245956
PUBMED 8482365
REFERENCE 2 (bases 5824 to 8669)
AUTHORS Bukreyev,A., Volchkov,V.E., Blinov,V.M. and Netesov,S.V.
TITLE The GP-protein of Marburg virus contains the region similar to the 'immunosuppressive domain' of oncogenic retrovirus p15E proteins
JOURNAL PNAS Lett. 323 (1-2), 183-187 (1993)
MEDLINE 93265932
PUBMED 8495737
REFERENCE 3 (bases 1 to 19112)
AUTHORS Bukreyev,A.A., Volchkov,V.E., Blinov,V.M., Dryga,S.A. and Netesov,S.V.
TITLE The complete nucleotide sequence of the Popp (1967) strain of Marburg virus: a comparison with the Musoke (1980) strain
JOURNAL Arch. Virol. 140 (9), 1589-1600 (1995)
MEDLINE 96028047
PUBMED 7487490
REFERENCE 4 (bases 1 to 19112)
AUTHORS Bukreyev,A.A.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-1994) Bukreyev A.A., Institute of Molecular Biology, Laboratory of molecular virology, Koltsovo, Novosibirsk region, Russia, 633159
FEATURES
source
1. .19112
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="Popp"
/db_xref="taxon:11269"
103..2190
/gene="NP"
103..2190
/gene="NP"
/functions="nucleoprotein"
/citations=[3]
/number=1
/codon_start=1
/label=NP
/product="NP protein"
/protein_id="CAA82536.1"
/db_xref="GI:450909"
/db_xref="GOA:P35263"
/db_xref="Uniprot/Swiss-Prot:P35263"
```



```
/translation="WDLHSLLELGTPTAPVHRNKKVILFTNTHQVSIICNOIIDAINS
GIDGLLEGGLLTLCVEHYNSDKOKFNTSPIAKYLRDAGYEFEDVWVKNADFLDV
IIPNPHYSPLTALKTLESTHYSQRIGLFSLFGLPFLKLVWGDRAISIEKALRQVTV
HQQOIVTPYNNMLTTHGMKVIIFGILRSSFTLKFVLIHQGVNLVTHGDAYIIISNV
GTEQFSGLLIIVKTLEFTLQKTDGCVTLHPLVRTSKVNEVASFPKALSNLARHGEVA
PFARVNLGGINNLEHGLYPOLSAIAGVATAGSTLAGVNVGVOYQOOLREAAHDAEI
KLQRHHEQEIQTAEDDEERKILEQHLQKTEITHSOTLAVLSQKREKLRLAAEIE
NNIVEDQOFKQNOVSFLNDPTFVEVTVQARPINRPTALPVPKQKLEKLRLAAEIE
SSSSFVLDNPFALLNDEDETLDDSVMIPTSTREFQGI PAPPQSQDLNNSQKQE
DESTNPKIQELPQELPPVOEDDESEYTTDSQESIQDQSGDNQGVLDLPPPIYAOE
KRODP IOHPAVSSODPRGSI GVDNGDILEPIRSPSSAPOEDTRAREAYELSPDFTN
YEDNQNPQVRVTKKRTFLYPNDLLQTNPPESLITAVEYQNPVSAKELQADWPD
NSFDERRHVAANL"
2944..3933
/gene="VP35"
2944..3933
/gene="VP35"
/function="minor component of RNA-dependent
RNA-polymerase?"
/citation=[1]
/citation=[3]
/number=2
/codon_start=1
/label=VP35
/product="VP35 protein"
/protein_id="CAA82537.1"
/db_xref="GI:450910"
/db_xref="UniProt/Swiss-Prot:Q03039"
/translation="WDSYSYQVQSEGLMTGKVPIDOVFGANPSEKHLHKBKPGTGVG
LQCSPLMSKATSTDIYVWDQLIVKTLADLLIPLNROIISDIQSTLNEVTRVHEIER
QAEITPVLKMGRTLEAISKGSEMLAKYDHLVISTGRTTAPAAAFADYLNHEHGVPPP
QPAIFPKLQVAGQACSKGTVMVNETTDAADKMSVLESEETFSKPNLSAKDALLIF
THLPNNTPFHILAQVLISKIAYKSGKGFADLPHQILSEGENAQAALTRLSTFADF
LGVPPVIRVKNFQTVPRPCKSLRAVPPNPTIDKGWCVYVSSEQGETRALKI"
4567..5478
/gene="VP40"
4567..5478
/gene="VP40"
/function="matrix protein?"
/citation=[1]
/citation=[3]
/number=3
/codon_start=1
/label=VP40
/product="VP40 protein"
/protein_id="CAA82538.1"
/db_xref="GI:450911"
/db_xref="GOA:P35254"
/db_xref="UniProt/Swiss-Prot:Q03040"
/translation="MASSSNTYNTQYLNPPPYADHGANQLIPADQLSNQQGITPNYV
GDNLDDQKGNVCHAFTEALEAIIIDISAYNEPTVKGPVPAWLPLGIMNSPEYPLAHTVAA
LLTGSYTTTQFTHNGQKFRVNRLTGIPAHPLRLMREGNOAFTQNMVI PRNFSTNQF
TYNLTNLVLSQLPDDAWRPSKDKLIGNTHMPAVSIHPNLPPIVLPTVKKQAYRQHK
NPNGPLAISGILHQLRVEKVPKTSLSRSLPADMFVREGMMKKRGNSPVVYFQ
APENPLNGFNRRQVLYANPTLSAV"
5940..7985
/gene="GP"
5940..7985
/gene="GP"
/function="envelope transmembrane glycoprotein"
/citation=[2]
/citation=[3]
/number=4
/codon_start=1
/label=GP
/product="GP protein"
/protein_id="CAA82539.1"
/db_xref="GI:450912"
/db_xref="GOA:P35254"
/db_xref="UniProt/Swiss-Prot:P35254"
/translation="WKTTCLESLILIOGIKTLPILEITASNNQPNQVDSVCSGTLOKT
EDVHLNGTILSQKQVADSPLEASKRAWAFRTGVPKKNVYTEGEERKACYNISVTDPSG
KSLLLDPPNTINDYPRCKRTIIHHIQGNPHQAQIALHLWGAFFLYDRIASTVYRGRVF
```

```
TEGNIAMIYVNTVHKMIFSRQGGYRHMNLSTNKKYWTSSNGTQTNDTCGFGALOXY
NSTKQTCASQPSPLPTARPEIKPTSTPTDATTINTDIPNDNDEDLITSGSGSGEQ
EPTKDAVTKQGLSMPPTSPSTPOQEGNNTDHSQGVTVPEPKNTLITTAQSPMP
PHNTAISTNNTSKNFSSTLSQNTNTDQSTATENEQTSAPSKTLPTPTGNLT
AKSTNNTGKPTTAPNNTNGHLTSPSTPNPTTQHLVYFRKRSILWREGMFPFLDG
LINAPIDFVPNTKTIIDESSSSGASAEEDQASPNISLTLSPFNINENTAYSGEN
ENCDALRLTWSQEDDLAAGLSWIPFPGIEGLYTAGLTKNONNLVCRRLRANOT
AKSLDLRLVTEERTFSLNRHAIDFLLRWGTCKVLGPDCCIGIEDLRLNISEQI
DQTKDQKREGTGWLGKRWKWTDMGVLTNLGLLLLSIAVLIALSCICRIFTKIYIG"
8867..9712
/gene="VP30"
8867..9712
/gene="VP30"
/citation=[3]
/number=5
/codon_start=1
/label=VP30
/product="VP30 protein"
/protein_id="CAA82540.1"
/db_xref="GI:450913"
/db_xref="GOA:P41326"
/db_xref="UniProt/Swiss-Prot:P41326"
/translation="MQOPRGRSTRNRHQTSASSIYHETQLPSKPHYTNHHPARMSST
RSSAESPTNHIIPRAPPTPNLSKPPPPKDMCMKMGIGLCTDPTCNRDHLDNLT
NRELLLMARKMLPNTDKTFRSLQDCGSPSLSKSGKQSQQTQKDVLTLENLHILNY
LHRSDIGKDETSRRAALSITCAGIRKTRSLINTMTLHINHENLPQDQGVIRKQTY
TGHLDKGGQFEAALWQWDKRSISLFVQAALYVMNIPCESSTSVQASYDHFILPQS
QSKGQ"
10205..10966
/gene="VP24"
10205..10966
/gene="VP24"
/citation=[3]
/number=6
/codon_start=1
/label=VP24
/product="VP24 protein"
/protein_id="CAA82541.1"
/db_xref="GI:450914"
/db_xref="UniProt/Swiss-Prot:P41325"
/translation="MAELSTRYNLPANVTEKSNLDLNSTARWIKEPSVGWTVKMGW
FVPHIINTGMALLHLKSNFVVPWQOTRNLFSHLFKNPKSTIIEPTLALRILLGVAL
KDQLQSLIPGFRSIVHMLSEWLLLETSRAIHSPNLLGLIYLSDMFKILMAGVNF
FNKMFTLHVNDHGKPSSEIKLTGQQIIITRVNMGFLVVRRIIDIEPCCGETVLSES
WVFLVAEAVLRHSEMQEKGQPLDTQYMSKIAI"
11479..18474
/gene="L"
11479..18474
/gene="L"
/function="major component of RNA-dependent
RNA-polymerase"
/citation=[3]
/number=7
/codon_start=1
/label=L
/product="L protein"
/protein_id="CAA82542.1"
/db_xref="GI:450915"
/db_xref="GOA:P35262"
/db_xref="UniProt/Swiss-Prot:P35262"
/translation="MQHPQYDARLSPILOQDCLLARSGLYSHYSHNPKLNCR
IPHHIYLRNSTALKTFLQNCISILTVFPHSISWHLTISIYDAINHVDVDFKYLPSSEL
VKYANWDNEFLKYLNLGLDHFVPAPASQWEDFPKERNPYMGSLQKGLRVMYRD
RIKQGRSLRSNWKPIGTDLELFGIADPIIFKVPVKTIIRNAVSLQASKPGRLVMYRD
ONLTPYLCDDDEFIVSAYSECFIMIKDVFIERYTWETICARAWLEDSDGADYLPDVL
GELYNQSDQIIAMYLEDGFKLILHLEPLCSCIOTHGIFTPGKYWFQSORLESYVEEL
CSLNKPKISGNKAECAQNFIKTLIQGKLTFFQCYCELSFLQKHGHPVLYIDVALDKV
KHAQSQKILPKPKYMFMTFCVFKIVAKNHYSGQSWYKTMIDHLTLTPYLQKHVSNIS
PPSQAEIYQHLWEWYFVEHEPLFTSTKISLSDISLFIKDRATVANQECWDSVDFRSVLGY
```

Query Match

72.0%; Score 1473.2; DB 14; Length 19112;

Best Local Similarity 82.5%; Pred. No. 0; Mismatches 358; Indels 0; Gaps 0; Matches 1688; Conservative 0;

QY 1 ATGAAGACCATATATTTTCTGATTTAGTCTCTCAATTTTATCCAAAGTATAAAACTCTCCCT 60
Db |||||
5940 ATGAAGACCATATGCTCTTTATCAGTCTTATCTTAATCCAAAGGATAAAAAATCTCTCCCT 5999
QY 61 GTTTTAAAGAAATTTGCTAGTAAACAGCCCACTCAAGATGTAGATTCAGTGTGCTCCGGAAAC 120
Db |||||
6000 ATTTTAAAGATAGCTAGTAAACATCAACCCCAAAATGTGAATTCGATGCTCCGGAACT 6059
QY 121 CTCCAAAGACAGAGATGTTTCTATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT 180
Db |||||
6060 CTCCAAAGACAGAGATGTTTCTATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT 6119
QY 181 GATTTCCCTTTTGAAGCATCTAAAGCATGGGCTTTTCAGGACAGGTGTCCTCCCAAGAAC 240
Db |||||
6120 GATTTCCCTTTTGAAGCATCTAAAGCATGGGCTTTTCAGGACAGGTGTAACCTCCCAAGAA 6179
QY 241 GTTGAGTATACGAAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT 300
Db |||||
6180 GTTGAGTATACGAAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT 6239
QY 301 TCTGAAAATCTTCTGCTCTGATCTCTCCAGTAAATATCCGCGATTAACCTAAATGTAAA 360
Db |||||
6240 TCTGAAAATCTTCTGCTCTGATCTCTCTACCAACATCCGCTGACTATCTTAATGCAA 6299
QY 361 ACTGTTCAATATTTCAAGGTCAAAACCTCATGCAAGGGGATTTGCCCTCCATTTGTGG 420
Db |||||
6300 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCGCAAGGGATCGCCCTCCATTTGTGG 6359
QY 421 GGGGCAATTTTCTGTATGATCGGTGCTCTCAACAATGTACCGAGGCAAGGCTTTC 480
Db |||||
6360 GGAGCAATTTTCTGTATGATCGGTGCTCTCTCAACAATGTACCGAGGCAAGGCTTTC 6419
QY 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATGAAGCAGTTTCAAGAAATGATTTTTCT 540
Db |||||
6420 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATGAAGCAGTTTCAAGAAATGATTTTTCT 6479
QY 541 AGGCAAGCAAGTTATTCGTCAACATGAACTTTGACCTCCACCAATAAATATTTGCAAGC 600
Db |||||
6480 AGGCAAGCAAGTTATTCGTCAACATGAACTTTGACCTCCACCAATAAATATTTGCAAGC 6539
QY 601 AGCAATGAACGCAAGAAATGATACGGGATTTTGGCATCTCTCCCAAGAAATACAACTCC 660
Db |||||
6540 AACAAATGAACCAAGCAACCAAGTAACTGATGCTTCCGCTCTCTCAAGAAATACAACTCC 6599
QY 661 ACAACCAATCAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCCAACAGTAACTCCG 720
Db |||||
6600 ACGAAGATCAACATGCTGCTCCGTTCCAAATACCTTCACTGCCCAAGCCGCTCCA 6659
QY 721 AGCAATTCATCTTACAAATATCTCAAAATTAATCTGCTAAATCTGGAACCTATGAACCCCAAGT 780
Db |||||
6660 GAGATCAAAACCAAGCAACCAAGTAACTGATGCTTCCGCTCTCTCAAGAAATACAACTCC 6719
QY 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAACAGGGGCCCCACACA 840
Db |||||
6720 AATGATGATGAGGACCTTATGAATCTTCCGGTTCCGGGTTCCGGGATCAAGGAAACCTTATACA 6779
QY 841 ACTCTTAATGATGATCACTGAACAGAAACAATGCTCAACAATATTTGTCACCTCTTCACTA 900
Db |||||
6780 ACTTCAGATGCGGTCACTTAAGCAAGGGGTTTTCATCAACATGCCACCCATCTCCCTCACA 6839
QY 901 CATCAAGCACTCTCAACATGAGCAAAACAGTACGAATTCCTTCCCGACATGCTGTAACT 960
Db |||||
6840 CAACCAAGCAGCCACAGCAAGAAAGAAACAACACAGACCATTTCCCAAGGTACTGTGACT 6899
QY 961 GAGCACAATGGAACCGACCAACAACAACACAGCAAGCGTCTCTCAACATTAATATACA 1020
Db |||||
6900 GAACCCAAACAAACCAACAACAACAGCAAGCGTCTCTCAACCGTCCATGCCCCCCCCCAACACTGCA 6959
QY 1021 ACTCCCACTTAACACTCTCAAGTACAACTCTCAGTACTCTCTTCCCTCCCAACCGCAAC 1080
Db |||||
6960 ATCTCTACTTAACACACTCTCAAGTACAACTCTCAGTACTCTCTCTGTATCACTACAAAC 7019

RESULT 12

AY358025

LOCUS

DEFINITION

Marburg virus strain M/S.Africa/Johannesburg/1975/ozolin, complete

19151 bp RNA linear

VRL 27-AUG-2003

QY 1081 ATCACCATTAATGATACACACGCTGAAGTACGAGAAAGCGAACAACCAATGCTCAGTTG 1140
Db |||||
7020 ACCACCAATTAACGACACACAGAGCACAGCCACTGAAATGAACCAACAGTGCCTCCCTCG 7079
QY 1141 AACACAATCTTAGATCCAAACAGAAAAATCCCAACACAGACAGCAAGACCAACAGACACAAC 1200
Db |||||
7080 AAAACAACCTGCTCTCCACAGGAAATCTTACCAACAGCAAGAGCACTTAACAACAGGAA 7139
QY 1201 AACATCATCATGACGACATCAGATATAACAGCAAAACACCCCAAAATTTCTTCCGGAT 1260
Db |||||
7140 GGGCCCAACCAACAGGCAACCAATATGACAAATGGGCAATTTAAACAGTCCCTCCCCACC 7199
QY 1261 TCTAGTCGCAACCCGCCCTCTATATCTTTTGAAGAAAGCAACGATCGATCTCTGGAGG 1320
Db |||||
7200 CCCAACCCGACACACAAATCTTGTATATTTTCAAGAAAGAAACGAAGTATCTCTCTGGAGG 7259
QY 1321 GAAGGACATGTTTCCCTTTCTGATGGGTTAATAATGCTCCAAATGATTTTGGACCA 1380
Db |||||
7260 GAAGGACATGTTTCTCTTCTGCAAGGTTAATAATGCTCCAAATGATTTGATCCA 7319
QY 1381 GTTCCAAATACAAAACAAATCTTTGATCAATCTCTAGTTCCTGCTGCTCGGCTGAGGAA 1440
Db |||||
7320 GTTCCAAATACAAAACAAATCTTTGATCAATCTTTAGTTCCTTAATTAATAAGTAA 7379
QY 1441 GATCAACATGCTCCGCCCAATATCAGTTTAACTTTATCTTTTCTTAATAATAAGTAA 1500
Db |||||
7380 GATCAACATGCTCCGCCCAATATCAGTTTAACTTTATCTTTTCTTAATAATAAGTAA 7439
QY 1501 AACACTGCTACTCTGAGGAAATGAGAAATGATTTGATGAGAGTAAAGAAATTTGGAGC 1560
Db |||||
7440 AACACTGCTACTCTGAGGAAATGAGAAATGAGAAATGATTTGATGAGAGTAAAGAAATTTGGAGC 7499
QY 1561 GTTCAGAGGATGACTGCGCGCAGGGCTCAGTTGGATACCGTTTCTGCTGCTCGGCTGGAAT 1620
Db |||||
7500 GTTCAGAGGATGACTGCGCGCAGGGCTCAGTTGGATACCGTTTCTGCTGCTCGGCTGGAAT 7559
QY 1621 GAAGGACTTTTATCTGCTGCTTAAATTAATAATCAAAACAAATTTGGTCTGCAAGTTGAG 1680
Db |||||
7560 GAAGGACTTTTATCTGCTGCTTAAATTAATAATCAAAACAAATTTGGTCTGCAAGTTGAG 7619
QY 1681 CGTCTAGCAATCAAACTGCCAAATCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
Db |||||
7620 CGTCTAGCAATCAAACTGCCAAATCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 7679
QY 1741 AGAATCTCTCTTAAATCAATAGACATGCTATTGATCTTCTACTCAAGATGAGGAGGA 1800
Db |||||
7680 AGGACTTTTCTTAAATTAATAGACATGCTATTGATCTTCTACTCACAAGTGGGGGGA 7739
QY 1801 ACATGCAAGTCTTGGACCTGATTTGTCATCGGATAGAGACTTGTCCAAAATATTT 1860
Db |||||
7740 ACATGCAAGTCTTGGACCTGATTTGTCATCGGATAGAGACTTGTCCAAAATATTT 7799
QY 1861 TCAGACAAATTTGACCAAAATTAAGAGCAACAAAGAGGGGAGCTGTTGGGGCTG 1920
Db |||||
7800 TCGGAAACAAATTTGACCAAAATTAAGAGCAACAAAGAGGGGAGCTGTTGGGGCTG 7859
QY 1921 GGTGTAATGCTTGGACATCCGACTGGGGTCTTCTTAACTTTGGGCAATTTTGTCTACTA 1980
Db |||||
7860 GGTGTAATGCTTGGACATCCGACTGGGGTCTTCTTAACTTTGGGCAATTTTGTCTACTA 7919
QY 1981 TTATCCATAGTCTTGAATGCTCTATCTCTGATATTTGCTGATCTTTTACTAAATATATC 2040
Db |||||
7920 TTATCCATAGTCTTGAATGCTCTATCTCTGATATTTGCTGATCTTTTACCATAATATATC 7979
QY 2041 GGATAA 2046
Db |||||
7980 GGTAA 7985

SVMGDNOCITLTLSPFPIDAPNDVQENAEELNARVAVELAITTCYSYGI FLKPEFTFVH
SGPIYFKKKYQLNGVQIPQSLKTKWARGCPISDSIPDDIQSGLASIGSFERSGTSRTH
IFPRWATASPHLMALINLLNQHILGFPILGFENIDISCPRKPLTFSEKUALIITIPQVLGG
LSFLNPKFLFYRNISDPLTSGFLGNALFELGKFLFYLIIAKKPGLDASDFVNP
LGLNVPGRREITITLQTVRENITITSONRIINSLFHIGSDLEQKYCEWLLSNPYM
SRFAADI FSRTPSKRLQVLGTQVTRTLAASRTISLTTEGTMKMLKTRNRRKVM
FSYIDALDDLSBSLEKFTCTVDVANTFLRAVSDVLKGRLLIGATILPCLLEQPKVM
INLESDIREOPNISSESELITNLLPYDCKELRLGGSHDTNLYVSCALDRKVVQKHS
VKRLAWTIGNRPIYIGRTEDKIGTFLPRVNSPAAKEATEMVSRLLWVTQGTADRE
KLLIPLNSRVNDIYQVNLPIFYVNSGNIVHRYNDQYQHSFWANRMSNTSTRAIIS
TNTLKYAGGQQAIDNII FQNTINGLAVADIALSLAKSSASNTFRMLMKNCT
RHVSEYLFDPKDLVDLNMKNBLVNDNDPLCSGIKGLGRVSRSTLSLSNVSDI
GSYDFPTAAWTIGETVGSIFSDSSQSTDPISSGCTKTFVTHFLVVPVIESIYFAG
ANLVESILSRLKISLNDLTLPLISSTIRNLSRSLRILLOSTFRHELVLTLAHHI
PLISLGGGAGEKSSDADFLFTASTQNFHNFSCMLKGGSSLPWLIFFPSEGOO
LKPILKILQTLSDLSLDPKQVCKTADTCCSDSFWYPSKSTRTHYINASTRLKIEY
KANKVQTPRSHLNCFLFLESSTSVSSNQVTSKYI VHPEDIPETINRTKLIRY
GSTALQMDIKWPPSEONVENCPSKDIRPKDNOKITKHDORVEKKESSOOQSPED
NMOTLA VMHNSPSQTTIKSIDVHEDFDASRVILNKNINFLNDCTINTNLLTPTPG
TEFLDTSPLOSSRYSSTPREBSLRSRQASVLYDCSNIPISLSDPFRMNSDONIQ
MLINAYKRDHLACHFDSQFCRTFQVSSMHYKGLYDLPPLPPELRKAI CLAGEGGGARL
LLKWKETDLHFFNTLATDSQOEABILSGRVIPRLMYNIDRLSALLESRRLINNLITQ
ITDITNPLWDSVLYQIPEPDSIDLTMDAETTKDREQLYKTI VNIWTRSPNPKIS
I IKVFLDYEGTFLMKNALQYGOVOLKKPYSSNAKNSWYLCSSKRRIORLOIDFP
DOVGIFLICKAMSRQORAI PMLKHIENKNYPASLHEPFLTLGPSPLESSCHRYTIPF
SEGKALPHKQVSYRQSKOHLHSLMLDYENNSPLLDLRNHFICSLRGIKTYNYNDILK
LNLVKAKEGKQSVLQVETLPPNHSVCI VHVHDECFCGERLLLLKLDIFIRNTKIAEQ
KLLNRVIGYILFFPFGLSKSRSLRA"
complement (19133. .19151)

primer_bind
ORIGIN

Query Match 70.5%; Score 1443.4; DB 14; Length 19151;
Best Local Similarity 81.6%; Pred. No. 0;
Matches 1669; Conservative 0; Mismatches 376; Indels 0; Gaps 0;

Qy	1	ATGAAGACCAATATTTTCTCATAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT	60
Db	5960	ATGAGACTACATGCTCTTATCAGTCTTATCTATCTATATCCAAAGGATATAAAACTCTCCCT	6019
Qy	61	GTTTATAGAAATGCTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTGCTCCGGAAAC	120
Db	6020	ATTTTGGAGATAGCAGTAACGATCAACCCCAAAATGGGATTCGGTATGCTCCGGAACT	6079
Qy	121	CTCCAAAGACAGAAAGATGTTCTATCTGATGGGATTTACATGATGGGCAAAAGATTGCT	180
Db	6080	CTCCAGAAAACAGAAAGATGTTCCATCTGATGGGATTTACATGAGCGGGCAGAAAATTGCT	6139
Qy	181	GATTCCCTTTTGGAGCATCTAAACGATGGGCTTTCAGGACAGGTGTTCTCCCAAGAAC	240
Db	6140	GATTCCCTTTTGGAGCATCTAAACGATGGGCTTTCAGGACAGGTGTTCTCCCAAGAAC	6199
Qy	241	GTTGAGTATACGGAAGGAGAGAGCCAAAACATGTTACATATAAGTGTAAACAGACCTT	300
Db	6200	GTTGAGTACACGGAAGGGAGAGCCAAAACATGCTACATATAAGTGTAAACAGATCCC	6259
Qy	301	TCTGGAATAATCTTGCTGCTGGATCTCCCAAGTAAATFCCCGGATATACCCTAAATGTAAA	360
Db	6260	TCTGGAATAATCTTGCTGCTGGATCTCCCAAGTAAATFCCCGGATATACCCTAAATGTAAA	6319
Qy	361	ACTGTTCAATCATATTCAGGTCAAAACCTCATGACAGGGGATTCGCCCTCATTTGTTGG	420
Db	6320	ACTATCCATCATATTCAGGTCAAAACCTCATGACAGGGGATTCGCCCTCATTTGTTGG	6379
Qy	421	GGGGCATTTTCTTGATGATCGGCTTGCCTCTACAAATGTATACCGGAGCAAGGTCTTC	480
Db	6380	GGAGCATTTTCTTGATGATCGGATTCGCCCTCCACAAATGTATACCGGAGCAAGGTCTTC	6439
Qy	481	ACTGAAGAAATATAGCAGCTATGATGTTAATAGACAGTTCACAGAAATGATTTTCT	540
Db	6440	ACTGAAGGAAACATAGCAGCTATGATGTTAATAGACAGTTCACAGAAATGATTTTCTCG	6499
Qy	541	AGGCAGACAGAGTTATCGTTCACATGAACCTTGACCTCCACCATAATATTTGGACAGC	600

Db	6500	AGGCAAGGACAAAGGTACCGTCAATGAATCTGACTTCTACTAATAAATATTGGACAAGT	6559
Qy	601	AGCAATGAAACGACAGAGAAATGATACGGGATGTTTTGGCATCTCCCAAGAAATACAATCC	660
Db	6560	AGCAAGGAAACGCAAAACAAATGACATCGGATGCTTGTGATCTCTTACGAGTACAAATCT	6619
Qy	661	ACAAACAAATCAAAACATCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG	720
Db	6620	ACGAAGAACCAAAACATGTGCTCCGTCTTAAACACCCCCACCCCGCCACAGCCGCTCGG	6679
Qy	721	AGCATTCACCTCTACAAATACTCAAAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	780
Db	6680	GAAATCAAAACCAACAGCACCCCAACCGATGCGCACTAGACTCAACACCAACACCAAAAC	6739
Qy	781	AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA	840
Db	6740	AGTGATGATGAGGATCTCTCAACATCCGGCTCGGGGTCTGGGGAAACAGGAACTCTATACG	6799
Qy	841	ACTCTTAATGTAGTCACTGAAACGAAACAACTCGTCAACAAATATTGTCACCTCTCTCACTA	900
Db	6800	ACTTCTGATGGGTCACTAAGCAAGGGCTTTCATCAACAAATGCGCACCCACTCCCTCACCG	6859
Qy	901	CATCCAAAGCACCTCAACACATGAGCAAAAACAGTACGAATCCTTCCCGACATGCTGTAACT	960
Db	6860	CAACGAGCAGCCACAGCAAGGAGGAAACAAACAAACCACTCCCAAGACGCTACAACCT	6919
Qy	961	GAGCAAAATGGAACCGACCCCAACAAACAAACAGCAACGCTCTCAACAAATTAATAACA	1020
Db	6920	GAACTTGATAACCAACAAACAACTGCAACCGCGCCACCGCTCTCCCAACACCAACCAACA	6979
Qy	1021	ACTCCCACTTATACATCTCAAGTACAACTCAGTACTCTCTCCCTCCCAACCGCAAC	1080
Db	6980	ATCTCCACCAACAAACCTCCAAACAAACCTCAGCAACCTCTCCCGACCAACCAAAAC	7039
Qy	1081	ATCACCAATATGATACAAACGCTGAACTAGCAGAAACGCAAAACCAATGCTCAGTTG	1140
Db	7040	ACCACCAATCCCAACACACAAAGCATGGCCACTGGAATGAGAAACCAAGTGCCTCCCGG	7099
Qy	1141	AACAACACTCTAGATCCAAACAGAAAAATCCCAACACAGAGCAAGACACCAACAGCAAC	1200
Db	7100	AAAGCAACCTGCTCCCAACAGAAAAATCTTACCAACAGAAAAAGACCAACAAATACAAA	7159
Qy	1201	AACATCATGACGACATCAGATATTAACAGCAAAACACCCCAACAAATCTTCTCCGAT	1260
Db	7160	AGCCCCACCAATGGAACCAATATACAAATGGAACATTTCTACTAGTCCCTCTCCAC	7219
Qy	1261	TCTAGTCCGACAAACCGCTCTCTATATCTTTAGAAAGAAACCATCGATCTCTGGAGG	1320
Db	7220	CCCAACTCGACTACAAACATCTTATATATTTAGGAGGAAACGAGATATCTCTGGAGG	7279
Qy	1321	GAAGGACATGTTCCCTTTCTGGATGGGTAAATAATGCTCCAAATGATTTTGACCCA	1380
Db	7280	GAAGGACATGTTCCCTTTCTAGATGGGTAAATAATGCTCCAAATGATTTTGATCCA	7339
Qy	1381	GTTTCAATACAAAAACAATCTTTGATGAATCTCTAGTTCGTGCTCCGCTGAGGAA	1440
Db	7340	GTTTCAATACAAAAACAATCTTTGATGAATCTCTAGTTCGTGCTCCGCTGAGGAA	7399
Qy	1441	GATCAACATGCTCCCAACAATATTAGTTTAACTTTTATCTTATTTCTTAAATAAATGAG	1500
Db	7400	GATCAACATGATCTCCCAACAATATTAGTTTAACTTTTATCTTATTTCTTCAACAAGTGA	7459
Qy	1501	AACACTGCCTACTCTGGAGAAAAATGAGAAATGATTTGATGACAGAGTTAAGAAATTTGAGC	1560
Db	7460	AACACTGCCTACTCTGGAGAAAAATGAGAAATGATTTGATGACAGAGTTAAGAAATTTGAGC	7519
Qy	1561	GTTTCAAGGAGATGATCTGGCGGAGGCTCAGTTGGATACCGTTTTTGGCCCTGGAAAT	1620
Db	7520	GTTTCAAGGAGATGATCTGGCGGAGGCTCAGTTGGATACCGTTTTTGGCCCTGGAAAT	7579
Qy	1621	GAAGGACTTTACATGCTGTTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	1680
Db	7580	GAAGGACTTTATACCGCAGGTTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	7639

QY 1681 CGCTAGCAATCAAACTGCCAACTCCTTGGAACTCTTATTGAGAGTCACAACTGAGGAA 1740
DB 7640 CGCTAGCAATCAAACTGCAAAATCCTTGGAACTCTTAAAGGTCACAAACCGAGGAA 7699
QY 1741 AGAACATTCCTCTTAATCAATAGACATGCTATTGACTTCTTACTACAAAGATGGGAGGA 1800
DB 7700 AGAACATTCCTTAATCAATAGACATGCTATTGACTTCTTACTACAAAGATGGGAGGA 7759
QY 1801 ACATGCAAGTCTTGGACCTGATTGTTGTCATCGGATAGAGACTTCTCCAAAAATATT 1860
DB 7760 ACATGCAAGTCTTGGACCTGATTGTTGTCATCGGATAGAGACTTCTCCAGAAATATT 7819
QY 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAAGAGGGGACTGGTTGGGGCTG 1920
DB 7820 TCAGAACAGATTGACCAAAATCAAGAGGACGACAAACAAAAAGAGGGGACTGGTTGGGGCTG 7879
QY 1921 GGTGTAATGTTGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTTGTCTACTA 1980
DB 7880 GGTGTAATGTTGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTTGTCTACTA 7939
QY 1981 TTATCCATAGTCTCTGATGCTCTATCCTGATTTCTGCTATCTTACTAAATATATC 2040
DB 7940 TTGTCATAGTCTCTGATGCTCTATCCTGATTTCTGCTATCTTACTAAATATATT 7999
QY 2041 GGATA 2045
DB 8000 GGATA 8004

RESULT 13
AF005733
LOCUS
DEFINITION
Marburg virus strain M/S.Africa/Johannesburg/1975/Ozolin
Glycoprotein precursor (GP) gene, complete cds.
ACCESSION
AF005733
VERSION
AF005733.1
KEYWORDS
GI:2459875
SOURCE
Lake Victoria marburgvirus
ORGANISM
Lake Victoria marburgvirus
Viruses; ssRNA negative-strand viruses; Mononegavirales;
Filoviridae; Marburg-like viruses.
REFERENCE
1 (bases 1 to 2046)
Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T., Bowen,M.D. and
Feldmann,H.
Auteurs
Variation in the glycoprotein and VP35 genes of Marburg virus
strains
JOURNAL
Virology 240 (1), 138-146 (1998)
MEDLINE
98110148
PUBMED
9448698
REFERENCE
2 (bases 1 to 2046)
Sanchez,A., Trappier,S.G., Stroher,U., Nichol,S.T. and Feldmann,H.
Direct Submission
Submitted (29-MAY-1997) Special Pathogens Branch, Centers for
Disease Control and Prevention, 1600 Clifton Road, Building 15,
Room S8611, Mail Stop G14, Atlanta, GA 30333, USA
TITLE
Location/Qualifiers
FEATURES
source
1..2046
/organism="Lake Victoria marburgvirus"
/mol_type="genomic RNA"
/strain="M/S.Africa/Johannesburg/1975/Ozolin"
/db_xref="taxon:11269"
1..2046
/gene="GP"
1..2046
/gene="GP"
/note="virion surface glycoprotein"
/codon_start=1
/product="glycoprotein precursor"
/protein_id="AAC40458.1"
/db_xref="GI:2459875"
EDVHLMGFTLSQKVADSPLEASKRWAFRTGVPPKNVYTBGEAKTCYNISVTDPSG

KSLLDPPTNRYDPKCTIHHIQGNPHAGGIALHLWGAFFLVRIASTTMYRGKVF
TEGNIAAMIYNKTVHMKIFSRQGGYRMLNTSTNKYNTNSNGTQTNDITGCLGLTLEY
NSTKNOTCAPSKTPPPPTARPEIKPTSTPTDRLNTNPNSDDEDLTSGSGSSEQ
ELTSDAVTKQGLSSTMPPTSPQCTPOGGNNTNHSODATTLELNDNTTADPTPT
SHNTTISTNNSTKHNLSLTSAPONTNNTOSMTVENEKTSAPPKATLPTPTPTT
EKSTNNTKSTTLEPNTINGHTSPSTPNSHTTOLLIYFRKRKRILILREGDMFPFLDG
LINAPIDFVPVNTKTIIDSSSSGASTEEDQHASNSILTSLSLXPLHTSGNTAYSGEN
ENDCDALRWSQVEDDLAAGLSWIFLPGFIEGLYTAGLTKNNQNLVCRRLRANOT
AKSLELLLRVTTEBRTPSLNRHAIDFLTRWGCTKVLGPDCTGIGEDLRLSEQI
DQIKDRQKQSGTGNGLGKWTSDWGLVTNLGILLLSIAVLIALSCICRIPTKYIG"
1..81
/gene="Gp"
82..2043
/gene="Gp"
/product="glycoprotein"
1930..2019
/gene="Gp"
/note="encodes transmembrane anchor peptide"
ORIGIN
Query Match 70.4%; Score 1441.2; DB 14; Length 2046;
Best Local Similarity 81.5%; Pred. No. 0;
Matches 1668; Conservative 0; Mismatches 378; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATATTTCTGATTAGTCTCAATTTTAAATCCAAAGTATAAAACTTCCT 60
DB 1 ATGAGGACTACATGCCTCTTTATCAGTCTTATCTTAATCCAAAGGATAAAACTTCCT 60
QY 61 GTTTTAGAATTGCTAGTAACAGCCAACTCAAGATGTAGATTGAGTGTCTCCGGAACC 120
DB 61 ATTTTGGAGATAGCCAGTAACGATCAACCCCAAAATGCGTATGCTCCGGAAT 120
QY 121 CTCCAAAGACAGAAAGATGTTCTATCTGATGGGATTTACATCGAGTGGGCAAAAAGTTGCT 180
DB 121 CTCCAAAAACAGAAAGATGTCATCTGATGGGATTTACATCGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTGGAAGCATCTAAACGATGGGCTTTCAGGACAGGTGTTCTCCCAAGAAC 240
DB 181 GATTCCCTTTGGAAGCATCTCAAGCATGGGCTTTCAGGACAGGTGTTCTCCCAAGAAC 240
QY 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACATATAGTGTACACACCTT 300
DB 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACATATAGTGTACACACCTT 300
QY 301 TCTGAAAAATCCTTGCTGCTGATCCTCCAGTAAATATCCCGCATTAATGCTTAAATGTA 360
DB 301 TCTGAAAAATCCTTGCTGCTGATCCTCCAGTAAATATCCCGCATTAATGCTTAAATGTA 360
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGCAAGGGATTTGCCCTCAATTTGG 420
DB 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCAAGGGATTTGCCCTCAATTTGG 420
QY 421 GGGGCAATTTTCTGTATGATCGGTTGCTCTACAAATGTACCGAGGCAAGTCTTC 480
DB 421 GGAGCAATTTTCTGTATGATCGGTTGCTCTACAAATGTACCGAGGCAAGTCTTC 480
QY 481 ACTGAAGGAAATATAGCAGCTATGATTGTTAATAGCAGTTTCACAGATGATTTTTTCT 540
DB 481 ACTGAAGGAAATATAGCAGCTATGATTGTTAATAGCAGTTTCACAGATGATTTTTTCT 540
QY 541 AGGCAAGGACAAAGTTATCGTCACATGAACCTTGACCTCCACCAATATAATATTTGGCAAGC 600
DB 541 AGGCAAGGACAAAGTTATCGTCACATGAACCTTGACCTCCACCAATATAATATTTGGCAAGC 600
QY 601 AGCAATGAACGACAGAAATGATCGGATGTTTGGCATCTCCCAAGNATACACCTCC 660
DB 601 AGCAATGAACGACAGAAATGATCGGATGTTTGGCATCTCCCAAGNATACACCTCC 660
QY 661 ACAACCAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAACCTCG 720
DB 661 ACAACCAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAACCTCG 720
QY 721 AGCATTCCTCTACAAATATCTCAAAATTAATCTGCTAAATCTGGAACTATGAACCCAAAGT 780

Db 721 GAATCAAAACCAAGACCCCAACCGATGCGACTAGACTCAACACCAACCAACCAAC 780
Qy 781 AGCGAGATGAGGACCTTTATGATTTTCGGGCTCAGGATCTGGAGAACAGGGGCCCAACACA 840
Db 781 AGTGATGATGAGGATCTCAACAATCCGGCTCGGGTCTGGGGAACAGGAACTCTATAGC 840
Qy 841 ACTCTTATGTAGTACACTGACAGACAGAAACAATCGTCAACAAATATGTCCACTCTCTCACTA 900
Db 841 ACTTCTGATGGGTCACTAAGCAAGGGCTTTTCATCAACAAATGCCACCCCACTCCCTCACCG 900
Qy 901 CATCAAGCACTCAACAATGAGCAAAACAGTACGAATCTTCCGACATGCTGTAACT 960
Db 901 CAACGAGCAGCCACAGCAGGAGGAGAAACAACAACCACTCCCAAGAGCGCTACAACT 960
Qy 961 GAGCACAATGGAACCGACCCCAACAACAACAGCAACGCTCTCTCAACAATACTAATAACA 1020
Db 961 GAACTTGATAACCAACAACAACAGTCAACCGCCAGCCCTCCCAACAACACCAACACA 1020
Qy 1021 ACTCCACCTATACACTCTCAAGTACAACTCTAGTACTCTTCCCTCCCAACCGGCAAC 1080
Db 1021 ATCTCCACCAACACCTCCCAACCAACCTCTAGCACCCTCTCCGACCAACCAACAAAC 1080
Qy 1081 ATCAACCAATATGATACACAACGCTGAATAGCAGAAAGCGCAACAACCAATGCTCAGTTG 1140
Db 1081 ACCACCAATCCCAACAACAAGCATGTCTACTGAAAATGAGAAACCAAGTGCCTCCCGC 1140
Qy 1141 AACCAACCTCTAGATCCCAAGAAAATCCCAACAGCAGACACCAACAGACACAAC 1200
Db 1141 AAGCAACCTGCCTCCCAACAGAAAATCTTACCAGAAAGAGCAACCAACAATACAAA 1200
Qy 1201 AACATCATGAGACATCAGATATAACAAGCAAAACACCCCAACAATTTCTTCGGAT 1260
Db 1201 AGCCCCACCAATGGAAACCAATACACAAATGACATTTCACTAGTCCCTCTCCACC 1260
Qy 1261 TCTAGTCGCAACCGCCCTCTATATACCTTATAGAAAGAACATCGATCCTCTGGAGG 1320
Db 1261 CCCAATCGACTACAAACATCTTATATATTTTCAGGAGGAACGAAGTATCCTCTGGAGG 1320
Qy 1321 GAAGCGCATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAAATTTGATTTGACCCA 1380
Db 1321 GAAGCGCATGTTCCCTTTCTAGATGGGTTAATAATGCTCCAAATTTGATTTGATCCA 1380
Qy 1381 GTTCCAATACAAAACAATCTTTGATGAATCTCTAGTCTGGTGCCTCGGCTGAGAA 1440
Db 1381 GTTCCAATACAAAACAATCTTTGATGAATCTCTAGTCTGGTGCCTTCAACTGAGAA 1440
Qy 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTTCTTAATATAATGAG 1500
Db 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATCTTATCTTCTCATACAAGTGA 1500
Qy 1501 AACACTGCTACTCTGGAGAAAATGAGAAATGATGTGATGACAGAGTTAAGAAATTTGGAGC 1560
Db 1501 AACACTGCTACTCTGGAGAAAATGAGAAATGATGTGATGACAGAGTTAAGAAATTTGGAGC 1560
Qy 1561 GTTCAGGAGATGACTCTGGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 1620
Db 1561 GTTCAGGAGATGACTCTGGCGCAGGGCTCAGTTGGATACCACTTTTGGCCCTGGAAATC 1620
Qy 1621 GAAGGACTTTACATGCTGTTTTTAATAATCAAAACAATTTGGTCTGCAAGTTGAGG 1680
Db 1621 GAAGGACTTTTACCGCAGGGTTTAAATTAATAATCAAAACAATTTGGTCTGCAAGTTGAGG 1680
Qy 1681 CGTCTAGCAATCAAACTGCCAAATCTTTTGAATCTTTTATGAGAGTCAACAATCTGAGAA 1740
Db 1681 CGTCTAGCAATCAAACTGCCAAATCTTTTGAATCTTTTATGAGGTCACAAACCGAGAA 1740
Qy 1741 AGAACAATCTCTTAATCAATAGACATGCTATTTGATCTTTCTACTCAACAAGTGGGGAGGA 1800
Db 1741 AGAACAATTTCTTAATCAATAGACATGCTATTTGATCTTTCTACTCAACAAGTGGGGAGGA 1800
Qy 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAAAGACTTGTCCAAAATATT 1860

Db 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGAATAGAAAGACTTGTCCAGAAATATT 1860
Qy 1861 TCAGAGCAAAATTGACCAAAATTTAAAGAGACGAAACAAAAGAGGGGACTGGTTGGGGCTG 1920
Db 1861 TCAGAACAGATTGACCAAAATCAAGAGAGACGAAACAAAAGAGGGGACTGGTTGGGGCTG 1920
Qy 1921 GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTTAACCTTTGGGCAATTTTGGTACTA 1980
Db 1921 GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTTAACCTTTGGGCAATTTTACTA 1980
Qy 1981 TTATCCATAGCTGCTTGTGATTTGCTCTATCCTGTATTTTGTCTATCTTTTACTAAATATATC 2040
Db 1981 TTGTCCATAGCTGCTTGTGATTTGCTCTATCCTGTATTTTGTCTATCTTTTACTAAATAT 2040
Qy 2041 GGATAA 2046
Db 2041 GGATAA 2046

RESULT 14
AX717717
LOCUS AX717717 2046 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 5 from Patent WO02079239.
ACCESSION AX717717
VERSION AX717717.1 GI:29890727
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Grogan, C.C.; Hevey, M.C. and Schmaljohn, A.L.
TITLE Chimeric filovirus glycoprotein
JOURNAL Patent: WO 02079239-A 5 10-OCT-2002;
U.S. Army Medical Research Institute of Infectious Diseases (US)
FEATURES
source
1..2046
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chimeric molecule between Marburg virus strain
Musoke Glycoprotein 1 and Marburg virus strain Raven
Glycoprotein 2"

ORIGIN
Query Match 64.5%; Score 1319.6; DB 6; Length 2046;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1592; Conservative 0; Mismatches 454; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAAACTCTCCCT 60
Db 1 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAATCTCCCC 60
Qy 61 GTTTTGAATTTGCTAGTACAGCCAACTCAAGATCTAGATTCACTGTGTCTCCGGAC 120
Db 61 ATTTTGAAGATAGTAGTAATAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAACT 120
Qy 121 CTCCAAAGACAGAAAGATGTTCTCATGTGGGATTTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCCAAAGACAGAAAGATGTTCTCATGTGGGATTTTACACTGAGTGGGCAAAAAGTTGCT 180
Qy 181 GATTCCCTTTGGAAAGCATCTAAACGATGGGCTTTTCAGGACAGAGTGTCTCCCAAGAAC 240
Db 181 GATTCCCTTTGGAGGATCCAGCGATGGGCTTTTCAGGACAGAGTGTCTCCCAAGAAC 240
Qy 241 GTTGAGTATACGGAAGAGAGCCAAACATGTTACATATATAGTGTACAGACCT 300
Db 241 GTTGAGTATACGAGGGGAGAGCCAAACATGTTACATATATAGTGTACAGACCT 300
Qy 301 TCTGAAAAATCTTGTCTGGATCTCTCCAGTAAATATCCCGATTATACCTTAAATGTAAA 360
Db 301 TCTGAAAAATCTTGTCTGGATCTCTCTACCAACATCCGTGACTATCCGAATGCAAA 360
Qy 361 ACTGTTTCATATATTCAGGTCAAAAACCTCTCATGACAGGGGATTTGCCCTCCATTTGTGG 420

361	Db		ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATCGGCCCTTCAATTTATGG	420
421	Qy	GGGGCATTTTTTCTTGTTATGATCGGGTTCCTCTACAAACAATGTATCCGAGGCAAGGCTTTC	480	
421	Db	GGAGCATTTTTTCTTGTTATGATCGGATTTGCCCTCCACAACAATGTATCCGAGGCAAGGCTTTC	480	
481	Qy	ACTGAAGGAATAATAGCAGCTATGATTTGTTAAATAAGACAGTTTCACAGAAATGATTTTTTCT	540	
481	Db	ACTGAAGGGAACATAGCAGCTATGATTTGTCAATAAGACAGTGCACAAAATGATTTTTCTCG	540	
541	Qy	AGGCAAGGACAAGGTTATTCGTCAATGAACTTGACCTTCCACCAATTAATAATTTGGACAGC	600	
541	Db	CGGCAAGGACAAGGGGTACCGTTCATATGAATCTGACTTCTAATAATAATTTGGACAAGT	600	
601	Qy	AGCAATGAACGCGAGAGAAATGATACGGGATGTTTTGGCATCTCTCCAAGAATAACAATCC	660	
601	Db	AGTAAACGGAAACGCAACGAATGACATCGGATGTTTTCGGCGCTCTTCAAGAAATACAATTCT	660	
661	Qy	ACAAACAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGGCCCAACAGTAATCCG	720	
661	Db	ACAAAGAACCAACATGTGCTCCGTCCAAATACCTCCACCACTGCCCAACAGCCCGTCCG	720	
721	Qy	AGCATTCACCTCTCAAAATACCTCAAAATTAATCTGTCTAAATCTGGAATCTATGAACCCAA	780	
721	Db	GAGATCAAACTCAACAGCACCCCAACTGATGCCACCACAAATCTCAATACACGGACCCCA	780	
781	Qy	AGCGACATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGGCCCCACACA	840	
781	Db	AGTGATGATGAGGACCTTCGCAACATCCGGCTCAGGGTCCGGAGAACAGAAACCCCAACA	840	
841	Qy	ACTCTTAATGTAGTCACTGAAACAGAAACAATTCGTCAACAATATTTGTCCACTCTCTTCACTA	900	
841	Db	ACTTCTGATGGGTTCACCAAGCAGGGCTTTCATCAACAATGCCACCCACTCCCTCACCA	900	
901	Qy	CATCCAAAGCACTTCACAACTAGACAAAAAGTACGAATCTTCTCCGACATGCTGTAACT	960	
901	Db	CAACCAAGCAGCCACAGCAGGAGGAGAAACAACAACCAATTCCTCCAAAGATGCTGTACT	960	
961	Qy	GAGCACAATGGACCGCCCAACAACCAACCAGCAACGCTCTCTCAACAATACTTAATACA	1020	
961	Db	GAACTAGACAAAAATACACAATCTGCACAACCTGTCATGGCCCCCTCATACACTACCA	1020	
1021	Qy	ACTCCCACTATAACACTCTCAAGTCAAACTCTAGTACTCTCTTCCCTCCAAACCCGCAAC	1080	
1021	Db	ATCTCTACTAACACACTCCAAACACACTTTCAGCACTCTCTCTGACCACTTACAAAC	1080	
1081	Qy	ATCACCAATTAATGATACAAACGTGAATAGCAGAGAAACGGAACCAATGCTTCAGTTG	1140	
1081	Db	ACCACCAATGACAAACACAGAGCACAACTCACTGAAAATGAGCAAAACGAGTGCCTCCCTCG	1140	
1141	Qy	AACACAACCTTAGATCCAAAGAAAATCCCAACACAGGACAAAGACACCAACAGCACACAC	1200	
1141	Db	ATAACAACCTTGCTCCAAACGGGAAATCCCAACACAGCAAAAGAGCACAGACAGCAAAA	1200	
1201	Qy	AACATCATATGAGCAGCATCAGATATAACAGCAAAACACCCCAAAATCTTCTCCGGAT	1260	
1201	Db	GGCCCCGGCACAAGCGCACCAACACAGCAAAATGAGCATTTTCAACAGTCTCTCCCCCACC	1260	
1261	Qy	TCTAGTCCGACAACCCCGCTCTATATACCTTTTGAAGAAACAGATCGATCTCTGGAGG	1320	
1261	Db	CCGAGCTCGACTGCACAACACTTTGTATATTTTCAGAAAGAAAGCGATCGATTTTCTGAAA	1320	
1321	Qy	GAAGGCGACATGTTCCCTTTTCTCGATGGGTTAATAATGCTCCAATTGATTTTGACCCA	1380	
1321	Db	GAAGGTGATATATTTCCGGTTTTTATAGATGGGTTAATAATACTGAAATTTGATTTGATCCA	1380	
1381	Qy	GTTCCAAATACAAAAACAACTTTTATGATGAATCCCTAGTTCTGGTGCTCGGCTGAGAAA	1440	
1381	Db	ATCCAAACACAGAAACAACTTTTATGATGAATCTCCGAGCTTTAATATTTCAACTAATGAG	1440	
1441	Qy	GATCAACATGCTCCCCCAATATATAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG	1500	

Db	1441	GAACAACACACTCCCGCGAATATCAGCTTTAACTTTTCTCTTATTTCTCTGATAAAATGGA	1501
Qy	1501	AACACTGCCCTACTCTCGAGAGAAATGAGAAATGATGTGATGCAGAGTTTAAGAAATTTGGAGC	1560
Db	1501	GATACTGCCCTACTCTCGGGGAAACGAGAAATGATTTGTATGTCAGAGTTGAGGATTTGGAGT	1560
Qy	1561	GTTTCAGGAGGATGACCTCGCCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTTGGAAATT	1620
Db	1561	GTGCAGGAGACGATTTGGCGGCGAGGGCTTAGCTGGATACCATTTTTTTGGCCCTTGGAAATC	1620
Qy	1621	GAAGGACTTTTACACTGCTGTTTTTAATTAATAATCAAAACAAATTTGGTCTCGAGGTTTCAGG	1680
Db	1621	GAAGGACTCTTACTGCCGGTTTAATCAAAATCAGAACAATTTAGTTTGTAGTTGAGG	1680
Qy	1681	CGTCTAGCCAAATCAAACTGCGAAATCCTTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA	1740
Db	1681	CGCTTAGCTAATCAAACTGCTAAATCCTTTGGAGCTCTTTTAAGGGTCAACAACGAGGAA	1740
Qy	1741	AGAACATTCCTTAAATCAATAGACATGCTATTTGACTTTCTACTCACAAGATGGGAGGA	1800
Db	1741	AGCACATTTTCTTAAATCAATAGGCATGCAATTTGACTTTTGTCTACGAGGTGGGCGGA	1800
Qy	1801	ACATGCAAAATGCTTGGACCTGATTTGTCATCGGATAGAAGACTTGTCCAAAAATATT	1860
Db	1801	ACATGCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAAGATCTATCTTAAAAATATC	1860
Qy	1861	TCAGAGCAAAATTCACCAAAATTAATAAGGACGAAACAAAAAGAGGGGACTGTTGGGGTCTG	1920
Db	1861	TCAGACAAATTCACCAAAATCAGAAAGATGACAAAAAGGAGGAACTGCTGGGCTCTA	1920
Qy	1921	GGTGGTAAATGTFGGACATCCGACTGGGGTGTTCTTACTTAACCTTGGGCATTTTGTCTACTA	1980
Db	1921	GGTGGCAAAATGGTGGACATCTGACTGGGGTGTTCTCACCAATTTGGGCATCTGCTACTA	1980
Qy	1981	TTATCCATAGCTCTCTGATTTGCTCTATCTCTGATTTGTCTGATCTTTACTTAATAATATC	2040
Db	1981	TTATCTATAGCTTTCTGATTTGCTCTGCTGTATCTGTCTGATCTTCTACTTAATAATATT	2040
Qy	2041	GGATAA 2046	
Db	2041	GGATGA 2046	
RESULT 15			
AX171715			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1. .1841			
/organism="synthetic construct"			
/mol_type="unassigned DNA"			
/db_xref="taxon:32630"			
/note="chimeric molecule between Marburg virus strain			
Musoke Glycoprotein 1 and Ebola virus Zaire Mayinga strain			
Glycoprotein 2"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 978; Conservative			
38.1%; Score 780; DB 6; Length 1841;			
74.8%; Pred. No. 7.4e-153;			
0; Mismatches 330; Indels			
0; Gaps			
0;			
Qy	1	ATGAAGACCATATATTTTCTGATTTAGTCTCATTTTAATCCAAAGTATAAAACCTCTCCCT	60

```

Db      1  ATGAAGACACATGTTTCCTTATCAGTCTTATCTTAAATTCAGGGACAAAAAATCTCCCC 60
Qy      61  GTTTTAGAAATTTGCTAGTAACAGCCAACTCAAGATGTAGATTTCAAGTGTGCTCCGGAAAC 120
Db      61  ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 120
Qy     121  CTCGAAAGACAGAAGATGTTCTATCTGTATGGATTTTACATGAGTGGGCAAAAAGTTGCT 180
Db     121  CTCGAAAGACAGAAGAGCTCCATCTGATGGGATTTCACTGAGTGGGCAAAAAGTTGCT 180
Qy     181  GATTCCCTTTGGAGCATCTAAACGATGGGCTTTCCAGACAGGTGTTCTCCCAAGAAC 240
Db     181  GATTCCCTTTGGAGCATCAAGGATGGGCTTTCCAGACAGGTGTTCTCCCAAGAAAT 240
Qy     241  GTTGTAGTATACGGAAGAGAGAGCCAAACATGTTACAATATAAGTGTAAACAGACCT 300
Db     241  GTTGTAGTACACAGAGGGGGAGAGCCAAACATGCTACATATAAGTGTAAACAGATCCC 300
Qy     301  TCTGAAATCTTGTCTGCTGATCTCCCAAGTAAATATCCGGATTAACCTTAATGTAAA 360
Db     301  TCTGAAATCTTGTCTGATCTCCCAAGTAAATATCCGGATTAACCTTAATGTAAA 360
Qy     361  ACTGTTTCATCATATTCAGGTCTAAACCTCTATGACAGGGGATGGCTCCATTTGTGG 420
Db     361  ACTATCCATCATATTCAGGTCTAAACCTCTATGACAGGGGATGGCTCCATTTATGG 420
Qy     421  GGGCATTTTCTGTATGATCGGTGCTCTACAAATGTATCCGAGGCAAGTCTTC 480
Db     421  GGAGCATTTTCTGTATGATCGGTGCTCTACAAATGTATCCGAGGCAAGTCTTC 480
Qy     481  ACTGAAGAAATATAGCAGCTATGATTTTAAATAGACAGTTTCAAGAAATGATTTTCT 540
Db     481  ACTGAAGGAAATATAGCAGCTATGATTTTAAATAGACAGTTTCAAGAAATGATTTTCT 540
Qy     541  AGGCAAGACAGGTATTCGTCAATGAACTTGACCTCCACCAATTAATATGACAAAC 600
Db     541  CGGCAAGGACAGGTATTCGTCAATGAACTTGACCTCCACCAATTAATATGACAAAC 600
Qy     601  AGCAATGAAACGACAGAAATGATACGGGATGTTTGGCTATCTCCCAAGAAATACAACTCC 660
Db     601  AGTAACGGAACGCAAGCAATGACACTGATGTTTGGGGCTCTTCAAGAAATACAACTCT 660
Qy     661  ACAACAAATCAAAATGCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAACCTCG 720
Db     661  ACAAGAAACCAACATGCTGCTCCGATCAAAATACCTCCACACTGCCCCACAGCTCCG 720
Qy     721  AGCATTTCACTCTAAGAAATCTCAATTAATTAATGCTGCTAAATCTGGAACTATGAA 780
Db     721  GAGATCAAACTCAACAGCACCCCACTGATGCAACCAAACTCAATPACCGAGCCCAAGC 780
Qy     781  AGCGAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db     781  AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACAGAACCCCAACA 840
Qy     841  ACTCTTAATGTAGTCACTGAACAGAAACAAATGCTCAACAAATATGTTCCACTCCTTCACTA 900
Db     841  ACTTCTGATGGCTCAACAGCAAGGGCTTTTCATCAACATGCCACCCCACTCCCTCACCA 900
Qy     901  CATCAAGACCTCAACATGAGCAAAACAGTAGCAATTCCTTCCCGACATGCTGTAACT 960
Db     901  CAACCAAGCACGCCACAGCAGGAGGAAACAAACAAACCATTTCCCAAGATGCTGTGACT 960
Qy     961  GAGCACAATGGAAACCGCCCAACACACACAGCAACGCTCTCAACAAATACTAATACA 1020
Db     961  GAACTAGCAAAAATAACAACTGCAACACCGCTCATGGCCCTCATAACTACCACTACCACA 1020
Qy    1021  ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCCGCAAC 1080
Db    1021  ATCTCTACTAACAACCTCCAAACAACTTACACTCTCTCTGACCATTAACAAC 1080
Qy    1081  ATCCAAATATGATACAAAGTGAATAGCAGAAAGGCAACAAACCAATGCTCAGTTG 1140

```

```

Db     1081  ACCACCAATGACAAACACACAGAGACCAATCACTGAAAATGAGCAAAACGAGTGCCTCC 1140
Qy     1141  AACACAACTCTAGATCCAAACAGAAAAATCCCAACAGGACAAAGACACCAACAGCAAC 1200
Db     1141  ATAAACACCTCTGCTCTCAACCGGAAATCCCAACAGCAACAGCAACAGCAACAGCA 1200
Qy     1201  AACATCATCATGACGACATCAGATATAACAGCAAAACACCCCAAAATTTCTTCCGGAT 1260
Db     1201  GGCCCCGCCCAACAGCGCACCAACACAGCAAAATGAGCATTTACACAGTCTCTCCCC 1260
Qy     1261  TCTAGTCCGACAAACCCCTCTCTATATCTTTTAAAGAAACGATCG 1308
Db     1261  CCCAGTCTGACTGCACAAACATCTTGTATATTTAGAAAGAGCGATCG 1308

```

Search completed: August 1, 2005, 02:43:54
Job time : 8799.18 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 13:55:23 ; Search time 1071.87 Seconds
(without alignments)

11299.693 Million cell updates/sec

Title: US-10-066-506A-7

Perfect score: 2046

Sequence: 1 atgaagaccatattttct.....ttactaaatatatcgataa 2046

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2002bs.*

8: geneseqn2003as.*

9: geneseqn2003bs.*

10: geneseqn2003cs.*

11: geneseqn2003ds.*

12: geneseqn2004as.*

13: geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2046	100.0	2046	10	ABT13453 DNA encod
2	1847.6	90.3	2046	10	ABT13456 DNA encod
3	1822	89.1	2247	4	AAD04043 Marburg v
4	1514.8	74.0	2103	8	ACC57764 Marburg v
5	1514.8	74.0	2164	4	AAD04042 Marburg v
6	1514.8	74.0	7778	10	ACC71550
7	1514.8	74.0	11460	3	AZ87211 VEE repli
8	1498.8	73.3	2100	3	AZ51039 Marburg v
9	1485.6	72.6	2051	10	ABT13454 DNA encod
10	1473.2	72.0	19112	13	ADM48345 Marburg v
11	1418.8	69.3	7005	10	ACC71551
12	1418.8	69.3	8256	10	ACC71552
13	1319.6	64.5	2046	10	ABT13452 DNA encod
14	834	40.8	6902	10	ACC71562
15	780	38.1	1841	10	ABT13451
16	746.6	36.5	2252	10	ABT13450 DNA encod
17	149	7.3	18890	13	ADM48327 Reaton eb
18	148.2	7.2	6885	10	ACC71538
19	148.2	7.2	6889	10	ACC71539
20	148.2	7.2	7001	2	AAx89795 DNA of pV

21	148.2	7.2	7002	10	ACC71536
22	148.2	7.2	7003	2	AAx59390
23	148.2	7.2	7023	10	ACC71541
24	148.2	7.2	7036	10	ACC71537
25	148.2	7.2	8146	10	ACC71540
26	145.8	7.1	2360	10	ADF28971
27	145.8	7.1	6913	10	ACC71531
28	145.8	7.1	8131	10	ACC71532
29	134.4	6.6	6940	10	ACC71535
30	134.4	6.6	7073	2	AAx59391
31	134.4	6.6	7082	10	ACC71533
32	134.4	6.6	7087	10	ACC71534
33	134.4	6.6	8221	10	ACC71547
34	134.4	6.6	8338	10	ACC71546
35	134.4	6.6	10783	10	ACC71545
36	134	6.5	2039	10	ABT13455
37	134	6.5	2172	4	AAD04041
38	134	6.5	2224	3	AZ51038
39	134	6.5	2298	3	AZ87189
40	134	6.5	2298	4	AAf76953
41	134	6.5	2298	12	ADL27460
42	134	6.5	6467	10	ACC71530
43	134	6.5	6561	10	ACC71524
44	134	6.5	6624	10	ACC71523
45	134	6.5	6914	10	ACC71529

ALIGNMENTS

RESULT 1

ABT13453
ID ABT13453 standard; DNA; 2046 BP.

XX ABT13453;

XX 30-JAN-2003 (first entry)

XX DNA encoding a chimeric filovirus protein RVN-GP1/MUS-GP2.

XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;

XX immune response; Ebola; Marburg virus; vaccine; gene; ds.

XX Unidentified.

XX Chimeric.

XX WO200279239-A2.

XX 10-OCT-2002.

XX 31-JAN-2002; 2002WO-US003339.

XX 31-JAN-2001; 2001US-0267522P.

XX (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.

XX Grogan CC, Hevey MC, Schmaljohn AL;

XX WPI: 2003-040651/03.

XX P-PSDB; ABJ18475.

XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2, useful for inducing an immune response against infection of different filoviruses, specifically against both Ebola and Marburg viruses.

XX Claim 18; Page 78-80; 94pp; English.

XX The invention relates to a chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2, where GP1 is from a filovirus different than that of GP2. The chimeric filovirus GP protein is useful for inducing an immune response against infection of different filoviruses, specifically against both Ebola and Marburg viruses by being used as a vaccine. This polynucleotide sequence represents a DNA encoding a chimeric filovirus

CC	protein of the invention	
XX		
SQ	Sequence 2046 BP; 662 A; 486 C; 379 G; 519 T; 0 U; 0 Other;	
	Query Match 100.0%; Score 2046; DB 10; Length 2046;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 2046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT 60	961 GAGCACAATGGAAACCGACCCCAACACACACACAGCAACGCTCTCCAAACATACTAATACA 1020
DB	1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT 60	1021 ACTCCACCTTATAACACTCTCAAGTACAACCTCTAGTACTCTCTCCCTCCCAACCGCAAC 1080
QY	61 GTTTTAGAAAATGCTAGTAACAGCCAACTCAAGATGTAGATTCACTGAGTGGGCAAAAGTTGCT 180	1021 ACTCCACCTTATAACACTCTCAAGTACAACCTCTAGTACTCTCTCCCTCCCAACCGCAAC 1080
DB	61 GTTTTAGAAAATGCTAGTAACAGCCAACTCAAGATGTAGATTCACTGAGTGGGCAAAAGTTGCT 180	1081 ATCAACCAATATGATACACACGCTGAACTAGCAGAAAAGCGAACAAACCAATGCTCAGTTG 1140
QY	121 CTCCAAGACAGAGATGTTTCATCTGATGGGTTTACACTGAGTGGGCAAAAGTTGCT 180	1141 AACCAACTCTAGATCCAAAGAAAATCCCAACACAGCAGCAGACACCAACAGCACAACC 1200
DB	121 CTCCAAGACAGAGATGTTTCATCTGATGGGTTTACACTGAGTGGGCAAAAGTTGCT 180	1141 AACCAACTCTAGATCCAAAGAAAATCCCAACACAGCAGCAGACACCAACAGCACAACC 1200
QY	181 GATTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTTTCTCCCAAGAAC 240	1201 AACCATCATGACGACATCAGATATAACAGCAAAACCCCAACCAAAATTTCTTCCCGAT 1260
DB	181 GATTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTTTCTCCCAAGAAC 240	1201 AACCATCATGACGACATCAGATATAACAGCAAAACCCCAACCAAAATTTCTTCCCGAT 1260
QY	241 GTTGAGTATACGGAAGGAGAGAACCCAAACATGTTTCAATATAAGTGTAAACAGCCCT 300	1261 TCTAGTCCGACAAACCCCTCTTATATATTTTAAAGAAACGATCGATCCTCTGGAGG 1320
DB	241 GTTGAGTATACGGAAGGAGAGAACCCAAACATGTTTCAATATAAGTGTAAACAGCCCT 300	1261 TCTAGTCCGACAAACCCCTCTTATATATTTTAAAGAAACGATCGATCCTCTGGAGG 1320
QY	301 TCTGGAATCTTCTGCTGATCCCTCCAGTAAATATCCGGATTTACCTTAATATGTA 360	1321 GAAGCGACATGTTTCCCTTTTCTGATGGGTTTAAATATGCTCCCAATTTGATTTGACCCA 1380
DB	301 TCTGGAATCTTCTGCTGATCCCTCCAGTAAATATCCGGATTTACCTTAATATGTA 360	1321 GAAGCGACATGTTTCCCTTTTCTGATGGGTTTAAATATGCTCCCAATTTGATTTGACCCA 1380
QY	361 ACTGTTTCATATATCAAGGTCAAAACCTCATGCAACAGGGGATGGCCCTCCATTGTGG 420	1381 GTTCCAAATACAAAACAACTCTTGTAGTGAATCTCTAGTCTGGTGCCTCCGGCTGAGGAA 1440
DB	361 ACTGTTTCATATATCAAGGTCAAAACCTCATGCAACAGGGGATGGCCCTCCATTGTGG 420	1381 GTTCCAAATACAAAACAACTCTTGTAGTGAATCTCTAGTCTGGTGCCTCCGGCTGAGGAA 1440
QY	421 GGGGCAATTTTCTGTATGATCGGCTCTCAACAAATGTACCGAGGCAAGGTCTTC 480	1441 GATCAACATGCTCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTAAATAAATGAG 1500
DB	421 GGGGCAATTTTCTGTATGATCGGCTCTCAACAAATGTACCGAGGCAAGGTCTTC 480	1441 GATCAACATGCTCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTAAATAAATGAG 1500
QY	481 ACTGAAGGAATATAGCAGCTATGATTTGTTAATAGACAGTTTACAGATGATTTTCT 540	1501 AACACTGCTACTCTGAGAAAATGAGAAATGATTTGTATGACAGAGTTAAGATTTGGAGC 1560
DB	481 ACTGAAGGAATATAGCAGCTATGATTTGTTAATAGACAGTTTACAGATGATTTTCT 540	1501 AACACTGCTACTCTGAGAAAATGAGAAATGATTTGTATGACAGAGTTAAGATTTGGAGC 1560
QY	541 AGGCAAGGACAGGTTTATCGTCACATGAATCTGACCTCCACCAATAAATTTGGACAAGC 600	1561 GTTCAGAGGAGTACCTGGCGCAGGCTCAGTTGGATACCGTTTGTGGCCCTGGAAT 1620
DB	541 AGGCAAGGACAGGTTTATCGTCACATGAATCTGACCTCCACCAATAAATTTGGACAAGC 600	1561 GTTCAGAGGAGTACCTGGCGCAGGCTCAGTTGGATACCGTTTGTGGCCCTGGAAT 1620
QY	601 AGCAATGAAACGACAGAAAATGATACGGGATGTTTGGCATCTCTCCAAAGAAATACA 660	1621 GAAGGACTTTTACACTGCTGTTTAAATTAATAATCAAAACAAATTTGGTCTGCAAGTTGAGG 1680
DB	601 AGCAATGAAACGACAGAAAATGATACGGGATGTTTGGCATCTCTCCAAAGAAATACA 660	1621 GAAGGACTTTTACACTGCTGTTTAAATTAATAATCAAAACAAATTTGGTCTGCAAGTTGAGG 1680
QY	661 ACAACATCAACATGCGCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAACCTCG 720	1681 CGTCTAGCCAATCAAACTGCCCCAAATCTTGGAACTCTTATTTGAGAGTCACTCACTGAGGAA 1740
DB	661 ACAACATCAACATGCGCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAACCTCG 720	1681 CGTCTAGCCAATCAAACTGCCCCAAATCTTGGAACTCTTATTTGAGAGTCACTCACTGAGGAA 1740
QY	721 AGCATTTCACTTACAAATACTCAAAATTAATCTGCTAAATCTGGAACATGAAACCAAGT 780	1741 AGAACATTTCTCTTAATCAATAGACATGCTATTTGACTTTCTACTCAACAGATGGGGAGA 1800
DB	721 AGCATTTCACTTACAAATACTCAAAATTAATCTGCTAAATCTGGAACATGAAACCAAGT 780	1741 AGAACATTTCTCTTAATCAATAGACATGCTATTTGACTTTCTACTCAACAGATGGGGAGA 1800
QY	781 AGCGACGATGAGGACCTTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840	1801 ACATGCAAAAGTCTGGACCTGATTTGTCATCGGATAGAGACTTGTCTCCAAATAAT 1860
DB	781 AGCGACGATGAGGACCTTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840	1801 ACATGCAAAAGTCTGGACCTGATTTGTCATCGGATAGAGACTTGTCTCCAAATAAT 1860
QY	841 ACTCTTAATGATGACCTGAAACAGAAAACATGCTCAACATATTTGTCACCTCTTCACTA 900	1861 TCAGAGCAAAATGACCAAAATTAAGAACGAAACAAAAGAGGGGAGCTGGTTGGGGCTG 1920
DB	841 ACTCTTAATGATGACCTGAAACAGAAAACATGCTCAACATATTTGTCACCTCTTCACTA 900	1861 TCAGAGCAAAATGACCAAAATTAAGAACGAAACAAAAGAGGGGAGCTGGTTGGGGCTG 1920
QY	901 CATCCAAGCAGCTCAACATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAAT 960	1921 GGTGTAATGTTGGACATCCGATGGGGTGTCTTACTACTACTGGGCATTTGGCTACTA 1980
DB	901 CATCCAAGCAGCTCAACATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAAT 960	1921 GGTGTAATGTTGGACATCCGATGGGGTGTCTTACTACTACTGGGCATTTGGCTACTA 1980
QY	961 GAGCAAAATGGAAACCGACCCCAACACACACAGCAACGCTCTCTCAACAACTAATACA 1020	1981 TTATCCATAGCTGCTTGAATTTGCTATCTCTGATTTGTCGATCTTTTACTAAATATATC 2040
		1981 TTATCCATAGCTGCTTGAATTTGCTATCTCTGATTTGTCGATCTTTTACTAAATATATC 2040
		2041 GGATAA 2046

Db	2041	GGATAA	2046					
RESULT 2								
ABT13456								
ID	ABT13456	standard; DNA; 2046	BP.					
XX								
AC	ABT13456;							
XX								
DT	30-JAN-2003	(first entry)						
XX								
DE	DNA encoding a chimeric filovirus protein RVN-GP1/GP2.							
XX								
KW	Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;							
XX	immune response; Ebola; Marburg virus; vaccine; gene; ds.							
OS	Unidentified.							
OS	Chimeric.							
XX								
PN	WO200279239-A2.							
XX								
PD	10-OCT-2002.							
XX								
PF	31-JAN-2002; 2002WO-US003339.							
XX								
PR	31-JAN-2001; 2001US-0267522P.							
XX								
PA	(USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.							
XX								
PI	Grogan CC, Hevey MC, Schmaljohn AL;							
XX								
DR	WPI; 2003-040651/03.							
XX	P-PSDB; ABJ18478.							
XX								
PT	New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,							
PT	useful for inducing an immune response against infection of different							
PT	filoviruses, specifically against both Ebola and Marburg viruses.							
XX								
PS	Disclosure; Page 89-91; 94pp; English.							
XX								
CC	The invention relates to a chimeric filovirus glycoprotein (GP) protein							
CC	comprising GP1 and GP2, where GP1 is from a filovirus different than that							
CC	of GP2. The chimeric filovirus GP protein is useful for inducing an							
CC	immune response against infection of different filoviruses, specifically							
CC	against both Ebola and Marburg viruses by being used as a vaccine. This							
CC	polynucleotide sequence represents a DNA encoding a chimeric filovirus							
CC	protein of the invention							
XX								
SQ	Sequence 2046 BP; 668 A; 481 C; 383 G; 514 T; 0 U; 0 Other;							
Query Match								
Best Local Similarity 90.3%; Score 1847.6; DB 10; Length 2046;								
Matches 1922; Conservative 0; Mismatches 124; Indels 0; Gaps 0								
Qy	1	ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT	60					
Db	1	ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT	60					
Qy	61	GTTTTAGAANTGCTAGTAAACGCCAACTCAAGATAGATTCACTGAGTGTCTCCGGAACC	120					
Db	61	GTTTTAGAANTGCTAGTAAACGCCAACTCAAGATAGATTCACTGAGTGTCTCCGGAACC	120					
Qy	121	CTCCAAAAGACAGAAGATGTTTCATCTGATGGATTTACACTGAGTGGCGAAAAGTTGCT	180					
Db	121	CTCCAAAAGACAGAAGATGTTTCATCTGATGGATTTACACTGAGTGGCGAAAAGTTGCT	180					
Qy	181	GATTCCCCCTTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGAGTGTTCTCTCCCAAGAAC	240					
Db	181	GATTCCCCCTTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGAGTGTTCTCTCCCAAGAAC	240					
Qy	241	GTTGAGTATACGGAAGGAGAGAAGACCCAAAAACATGTTTCAATATAGTGTAAAGACCCCT	300					
Db	241	GTTGAGTATACGGAAGGAGAGAAGACCCAAAAACATGTTTCAATATAGTGTAAAGACCCCT	300					


```
QY 1381 GTTCCAAATACAAAACAACTCTTTGATGATCCTCTAGTCTGTGTCCTCGGCTGAGGAA 1440
Db 1381 ATCCCAAAACACAGAAACAACTCTTTGATGATCCTCGGCTTTAATACCTCAACTAATGAG 1440
QY 1441 GATCAACATGCTCCGCCCAATATAGTTTAACTTTATCTTATTTCTTAAATATAATGAG 1500
Db 1441 GAACAACACACTCCGCCCAATATAGTTTAACTTTATCTTATTTCTTAAATATAATGAG 1500
QY 1501 AACACTGCTACTCTGGAGAAAATGAGAAATGATTTGATGATGATGATGATGATGATGAGC 1560
Db 1501 GATACTGCTACTCTGGGGAACACAGAAATGATTTGATGATGATGATGATGATGATGAGT 1560
QY 1561 GTTCAGGAGATGACTCGGCCAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db 1561 GTGCAGGAGGACGATTTGGCGGAGGCTTAGCTGATGATGATGATGATGATGATGATGATGAT 1620
QY 1621 GNAGGACCTTACACTGCTGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
Db 1621 GAAGGACTCTATACCTGCGGTTTAAATCAAAAATCAGAAATTTAGTTTGTAGTTGAGG 1680
QY 1681 CCTCTAGCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTAAATCCTTGGAGCTCTTGTAAAGGTCACAACTGAGGAA 1740
QY 1741 AGAACATTTCTTAAATCAATAGACATGCTATTTGATGATGATGATGATGATGATGATGATGAT 1800
Db 1741 AGGACATTTCTTAAATCAATAGGATGCAATTTGATGATGATGATGATGATGATGATGATGAT 1800
QY 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATGCGGATAGAGACTTCTCCAAAATATT 1860
Db 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATGAGAAATAGAGATCTATCTAAAAATATC 1860
QY 1861 TCAGAGCAAAATGACCAAAATAAAAAGGACGAAACAAAAGAGGAGCTGTTGGGCTCTG 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGATGAAACAAAAGGAGGAACTGGCTGGGCTCTA 1920
QY 1921 GGTGTAATGTTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTGCTACTA 1980
Db 1921 GGTGCAAAATGTTGGACATCTGACTGGGGTGTCTTACTAACTTGGGCAATCTGCTACTA 1980
QY 1981 TTATCATAGCTGCTTGAATGCTATCTGATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db 1981 TTATCATAGCTGCTTGAATGCTATCTGATGATGATGATGATGATGATGATGATGATGATGAT 2040
QY 2041 GGATAA 2046
Db 2041 GGATGA 2046

RESULT 3
ID AAD04043
XX AAD04043 standard; DNA; 2247 BP.
AC AAD04043;
XX
XX
DT 02-JUL-2001 (first entry)
DE Marburg virus Ravn strain glycoprotein (GP) DNA.
XX
KW Glycoprotein; GP; immune response; vaccine; antiviral;
KW type I transmembrane protein; ds.
XX
OS Marburg virus.
XX
FH Key Location/Qualifiers
FT CDS 97..2142
FT /*tag= a
FT /product= "Marburg virus Ravn strain glycoprotein (GP)"
XX
XX US200959-B1.
XX
PD 13-MAR-2001.
XX
```

```
PF 04-DEC-1996; 96US-00760615.
XX
PR 04-DEC-1996; 96US-00760615.
XX
PA (POWD-) POWDERJECT VACCINES INC.
XX
PI Haynes JR, Schmaljohn CS, Fuller DL, Schmaljohn A, Jahrling PB;
XX P-PSDB; AAE00708.
XX
XX WPI; 2001-280564/29.
XX
XX Inducing an immune response to a Marburg or Ebola virus involves
XX delivering a genetic vaccine, which contains a genetic construct encoding
XX antigenic determinants for filovirus, using a particle acceleration
XX device.
XX
XX Claim 6; Col 33-40; 33pp; English.
XX
XX The patent discloses a method of inducing an immune response to Marburg
XX or Ebola virus glycoprotein which involves delivering a genetic vaccine,
XX containing a genetic construct encoding antigenic determinants for
XX filovirus, using a particle acceleration device. The genetic vaccine for
XX filovirus is created by joining a DNA sequence encoding at least a
XX portion of the filovirus glycoprotein to a promoter effective to promote
XX transcription of the DNA sequence. This method is useful for inducing
XX humoral, cell-mediated and secretory immune responses in the treated
XX individual. The present sequence is a DNA encoding Marburg virus Ravn
XX strain glycoprotein (GP), which is a type I transmembrane protein. This
XX sequence is used in the construction of genetic vaccine against
XX filoviruses
XX
XX Sequence 2247 BP; 734 A; 518 C; 424 G; 571 T; 0 U; 0 Other;
XX
XX Query Match 89.1%; Score 1822; DB 4; Length 2247;
XX Best Local Similarity 93.2%; Pred. No. 0;
XX Matches 1906; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAACTCTCCCT 60
Db 97 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAACTCTCCCT 156
QY 61 GTTTTAGAATTTGCTAGTAACAGCAACCTCAAGATGATGATGATGATGATGATGATGATGATGAT 120
Db 157 GTTTTAGAATTTGCTAGTAACAGCAACCTCAAGATGATGATGATGATGATGATGATGATGATGAT 216
QY 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGATTTACACTGAGTGGCAAAAAGTGTCT 180
Db 217 CTCCAAAGACAGAAAGATGTTTCATCTGATGGATTTACACTGAGTGGCAAAAAGTGTCT 276
QY 181 GATTCCCTTTTGGAAAGCATCTAAACGATGGCTTTTCAGGACAGGTTTCTCCCAAGAAC 240
Db 277 GATTCCCTTTTGGAAAGCATCTAAACGATGGCTTTTCAGGACAGGTTTCTCCCAAGAAC 336
QY 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACATATATAGTGTACAGACCT 300
Db 337 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACATATATAGTGTACAGACCT 396
QY 301 TCTGAAAAATCCTTGTGCTGATCCTCCAGTAAATATCCCGGATTTACCTAAATGTAAA 360
Db 397 TCTGAAAAATCCTTGTGCTGATCCTCCAGTAAATATCCCGGATTTACCTAAATGTAAA 456
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAAACCTCATGACAGGGGATTTGCCCTCATTTGTGG 420
Db 457 ACTGTTTCATCATATTTCAAGGTCAAAAACCTCATGACAGGGGATTTGCCCTCATTTGTGG 516
QY 421 GGGGATTTTCTTGTATGATCGGTTGCTCTACACAATGTACCGAGGCAAGTCTTTC 480
Db 517 GGGGATTTTCTTGTATGATCGGTTGCTCTACACAATGTACCGAGGCAAGTCTTTC 576
QY 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAGACAGTTTCACAGAAATGATTTTTTCT 540
Db 577 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAGACAGTTTCACAGAAATGATTTTTTCT 636
```

QY 541 AGGCAAGGACAGGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTTGGACAGC 600
DB 637 AGGCAGAGCAGGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTTGGACAGC 696
QY 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTTGGCATCTCCCAAGAAATACAACTCC 660
DB 697 AGCAATGAACCGCAGAGAAATGATAGGGGATGTTTTGGCATCTCCCAAGAAATACAACTCC 756
QY 661 ACAACAAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACCTCG 720
DB 757 ACAACAAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACCTCG 816
QY 721 AGCATTTCACTTACAAATPACTCAAAATTAATCTGTAATCTGGAACCTATGAAACCCCAAGT 780
DB 817 AGCATTTCACTTACAAATPACTCAAAATTAATCTGTAATCTGGAACCTATGAAACCCCAAGT 876
QY 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
DB 877 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 936
QY 841 ACTCTTAATGATGACATGAAACAGAAACAAATCGTCAACAAATATTTGTCACACTCTCTCACTA 900
DB 937 ACTCTTAATGATGACATGAAACAGAAACAAATCGTCAACAAATATTTGTCACACTCTCTCACTA 996
QY 901 CATCCAGCACCTCACACATGAGCAAAACAGTACGATCTCTCCGACATGCTGTAAC 960
DB 997 CATCCAGCACCTCACACATGAGCAAAACAGTACGATCTCTCCGACATGCTGTAAC 1056
QY 961 GAGCACAATGGAACCGACCCCAACACACACAGCAACGCTCTCAACAAATATTAATACA 1020
DB 1057 GAGCACAATGGAACCGACCCCAACACACACAGCAACGCTCTCAACAAATATTAATACA 1116
QY 1021 ACTCCACCTATAACACTCTCAAGTACAACTCTGATCTCTTCCCTCCCAACCGCAAC 1080
DB 1117 ACTCCACCTATAACACTCTCAAGTACAACTCTGATCTCTTCCCTCCCAACCGCAAC 1176
QY 1081 ATCCACCAATATGATACACACAGTGAACCTAGCAGAAAGCGAACAACCAATGCTCAGTTG 1140
DB 1177 ATCCACCAATATGATACACACAGTGAACCTAGCAGAAAGCGAACAACCAATGCTCAGTTG 1236
QY 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACAGAGCAAGACACCAACAGCAACAC 1200
DB 1237 AACCAACTCTAGATCCCAACAGAAATCCCAACAGAGCAAGACACCAACAGCAACAC 1296
QY 1201 AACCATCATGACGACATCAGATATATAAGAGCAAAACACCAAAATTTCTTCOGGAT 1260
DB 1297 AACCATCATGACGACATCAGATATATAAGAGCAAAACACCAAAATTTCTTCOGGAT 1356
QY 1261 TCTAGTCCGACAAACCGCGCTCTTATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
DB 1357 TCTAGTCCGACAAACCGCGCTCTTATATCTTTAGAAAGAAACGATCTCTGGAGG 1416
QY 1321 GAAGCGACATGTTCCCTTTCTGATGGGTTAATAATGCTCCAAATGATTTTGACCCA 1380
DB 1417 GAAGGTGATATATTTCCCGTTTATAGATGGGTTAATAATGATGAAATGATTTGATCCA 1476
QY 1381 GTTCCAAATACAAACAAATCTTTGATGAATCTCTAGTTCTGCTGCTCGGCTGAGGAA 1440
DB 1477 ATCCCAACACAGAAACAAATCTTTGATGAATCTCCACGCTTAAATCTCAACTAATAG 1536
QY 1441 GATCAACATGCTCCCGCAGGCTCAGTTGATACCGTTTATCTTAATTTTCTTAATATAAATAG 1500
DB 1537 GAAACAACACACTCCCGCAGATCAGTTTAACTTTCTCTTATTTCTGATGAAAAATGGA 1596
QY 1501 AACACTGCTACTCTGGAGAAATGAGAAATGATGATGATGAGAGTTAAGAAATTTGGAGC 1560
DB 1597 GATCTGCTACTCTGGGGAACACAGAAATGATGATGATGAGAGTTGAGGATTTGGAGT 1656
QY 1561 GTTTCAGGAGGATGACTCGCGCAGGCTCAGTTGATACCGTTTATCTTAATTTTGGCCCTGGAAT 1620
DB 1657 GTGAGGAGGAGGATTTGGCGCAGGCTTAGCTGGATACCAATTTTGGCCCTGGAATC 1716
QY 1621 GAAGGACTTTACATGCTGTTTTTAATTAATAAATCAAAACAAATTTGGTCTGCAGGTTGAGG 1680

DB 1717 GAAGGACTCTACTGCGGTTTAAATCAAAATCAAGAACTTTAGTTTGTAGTTGAGG 1776
QY 1681 CGTCTAGCCAAATCAAACTGCCAAATCTTGAACTCTTAATGAGAGTCAAACTGAGGAA 1740
DB 1777 CGCTTAGCTAATCAAACTGCTAAATCTTGAGAGCTCTTGTAAAGGTCACAACCGAGGAA 1836
QY 1741 AGAACATCTCTTAATCAATAGACATGCTATTGACTTTCTACTCAACAGATGGGAGGA 1800
DB 1837 AGGACATTTTCTTAAATCAATAGCATGCAATTTGCTTACGAGGTGGGCGGA 1896
QY 1801 ACATGCAAGTGTCTTGACCTGATTGTCATCGGATAGAGACTTTGTCCAAAAATATT 1860
DB 1897 ACATGCAAGTGTCTAGACCTGATTGTCATAGGATAGAGATCTATCTAAAAATATC 1956
QY 1861 TCAGAGCAAAATTCACCAATTTAAAAAGGACGAAACAAAAAGGGGACTGTTTGGGCTG 1920
DB 1957 TCAGACAAATTCGACAAATTCAGAAAGGATGAAACAAAAAGGAGGAACTGGCTGGGCTA 2016
QY 1921 GGTGGTAAATGGTGACATCGGACTGGGTGTTCTTACTAACTTTGGGCAATTTTGTACTA 1980
DB 2017 GGTGGCAATTTGGTGACATCTGACTGGGCTGTTCTCACCATTTTGGGCACTCTCTACTA 2076
QY 1981 TTATCCATAGCTGCTTGTGATTGCTCTATCTGTATTTGTGCTATCTTTACTAAATATATC 2040
DB 2077 TTATCTATAGCTGTTCTGATTGCTCTGTCTGCTGATCTGCTGCTATCTTCACTAAATATATC 2136
QY 2041 GGATAA 2046
DB 2137 GGATGA 2142

RESULT 4
ACCS57764.
ID ACCS57764 standard; cDNA; 2103 BP.
XX ACCS57764;
AC ACCS57764;
XX 28-JUL-2003 (first entry)
XX Marburg virus envelope glycoprotein coding sequence.
DE Marburg virus envelope glycoprotein coding sequence.
XX Lenti-virus; pseudotyped virus; gene therapy; vector; gene; ss.
XX Marburg virus.
OS
XX Key Location/Qualifiers
XX CDS 10..2055
XX /*tag= a
XX /product= "Envelope glycoprotein"
XX WO2003035849-A2.
XX
XX 01-MAY-2003.
XX
XX 28-OCT-2002; 2002WO-US034545.
XX
XX 26-OCT-2001; 2001US-0353221P.
XX 26-OCT-2001; 2001US-0356436P.
XX
XX (MCCR/) MCCRAY P B.
XX (SAND/) SANDERS D A.
XX (DAVI/) DAVIDSON B L.
XX
XX Mccray PB, Sanders DA, Davidson BL;
XX
XX WPI; 2003-421416/39.
XX P-ESDB; ABR42242.
XX
XX New pseudotyped lentivirus comprising a lentiviral capsid, a lipid bilayer and a Marburg glycoprotein disposed in the lipid bilayer, useful for eliciting an immune response against feline immunodeficiency virus infection.

The present sequence is the coding sequence for the envelope glycoprotein of Marburg virus. The invention provides methods for gene transfer to cells using glycoprotein-pseudotyped lentiviruses. The glycoprotein is preferably Marburg glycoprotein. The Marburg viral glycoprotein preferably has a mutation in the C-terminal portion (see AB842243-47) that results in a higher titre production of the pseudotyped virus. The pseudotyped lentivirus is used to introduce nucleic acid sequences encoding a desired protein into a hepatocyte, brain glial or airway epithelial cell, in vitro or in vivo. The desired protein is preferably cystic fibrosis transmembrane conductance regulator, or is the low density lipoprotein receptor, alpha1-antitrypsin, ornithine transcarbamylase, Factor VIII or a high affinity glutamate receptor (all claimed)

XX	Sequence 2103 BP; 671 A; 533 C; 408 G; 491 T; 0 U; 0 Other;	
PS	Query Match	74.0%; Score 1514.8; DB 8; Length 2103;
XX	Best Local Similarity	83.8%; Pred. No. 0;
CC	Matches 1714; Conservative	0; Mismatches 332; Indels 0; Gaps 0

QY	1	ATGAAGACCATATATTTCTGATAGTCTCATTTTAAATCCAAAGTATFAAAAATCTCCCT	60
DB	10	ATGAAGACCATGTTTCTTATCAGTCTTATCTAATTCAAGGACAAAAATCTCCCC	69
QY	61	GTTTTAGAATGTCTAGTAACAGCCAACTCAAGATGTAGATTTCAGTGTCTCCGGAACC	120
DB	70	ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGGGATTCGGTATGCTCCGGAAT	129
QY	121	CTCCAAAGACAGAAGATGTTTCATCTGATGGGATTTACATCGAGTGGGCAAAAAGTTCCT	180
DB	130	CTCAGAAGACAGAAGACGTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTCCT	189
QY	181	GATTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC	240
DB	190	GATTCCCTTTTGAAGCATCTCAAGCGATGGGCTTTTCAGGACAGGTGTACCTCCCAAGAA	249
QY	241	GTTGAGTATACGGAAGNAGAGCCAAACATGTTACATATAGGTGTAACAGACCT	300
DB	250	GTTGAGTACACAGAGGGGGAGAGCCAAACATGCTACAATATAGGTGTAAACGATCCC	309
QY	301	TCTGAAAAATCCTTGCTCGATTCCTCCAGTAAATATCCGCGATTACCTAAATGTAAA	360
DB	310	TCTGAAAAATCCTTGCTGTAGATCCTCTACCAACATCCGTGACTATCCTTAATGCAAA	369
QY	361	ACTGTTTCATCATATTTCAAGGTCAAAAACCTCATGACACAGGGGATGTCCTCATTTGTG	420
DB	370	ACTATCCATCATATTTCAAGGTCAAAAACCTCATGACACAGGGGATGTCCTCATTTATGG	429
QY	421	GGGCGATTTTCTGTATGATCGGTGCTCTACCAAAATGTACCGAGGCAAGGTCTTC	480
DB	430	GGAGCATTTTCTGTATGATCGGATTCCTCCCAACAAATGTACCGAGGCAAGGTCTTC	489
QY	481	ACTCAAGGAAATATAGCAGCTATGATCTGTTAATAGACAGTTCACAGAAATGATTTTTTCT	540
DB	490	ACTGAGGGAAACATAGCAGCTATGATCTGTTAATAGACAGTTCACAGAAATGATTTTTCTG	549
QY	541	AGGCAAGGACAAAGTTATTCGTCACATGAATCTTGACCTCCACCAATATAATTTGGACAAGC	600
DB	550	CGGCAAGGACAAAGGTACCGTCATATGAATCTGACTTCTACTAATAAATATTGGACAAGT	609
QY	601	AGCAATGAAACGACAGAAATGATACGGGATGTTTTGGCATCTCCCAAGATATCAAACTCC	660
DB	610	AGTAACGGAAACGCAAAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT	669
QY	661	ACAAACAAATCAAAATGCGCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG	720
DB	670	ACAAAGAACCAAAATGTGCTCCGTCGTAATACTCTCCACACTGCCCCACAGCCGCTCG	729
QY	721	AGCATTTCACTACAAATFACTCAAAATTAATACTGCTAAATCTGGAACTATGAACCCAAAGT	780

QY 1861 TCAGAGCAAAATTGACCAATTAATAAGGACGACAAAGAGGGGACTGGTTGGGCTCG 1920
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1929
QY 1921 GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCATTTTCTCTACTA 1980
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1989
QY 1981 TTATCCATAGCTGCTTGGATGCTATCCCTGATATTTGCTGATCTTTACTAAATATATC 2040
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 2049
QY 2041 GGATAA 2046
DB ||||||| 2050 GGATAA 2055

RESULT 5
AAD04042
ID AAD04042 standard; DNA; 2164 BP.
XX
AC AAD04042;
DT 02-JUL-2001 (first entry)
XX
DE Marburg virus Musoke strain glycoprotein (GP) DNA.
XX
KW Glycoprotein; GP; immune response; vaccine; antiviral;
KW type I transmembrane protein; ds.
XX
OS Marburg virus.

Key Location/Qualifiers
FH 119..2164
FT CDS /*tag= a
FT /*product= "Marburg virus Musoke strain glycoprotein (GP)"
XX

US6200959-B1.
XX
XX 13-MAR-2001.
XX
XX 04-DEC-1996; 96US-00760615.
XX
XX 04-DEC-1996; 96US-00760615.
XX
XX (POWD-) POWDERJECT VACCINES INC.
XX
XX Haynes JR, Schmaljohn CS, Fuller DL, Schmaljohn A, Jahrling PB;
XX
XX WPI; 2001-280564/29.
DR P-PSDB; AAE00707.

Inducing an immune response to a Marburg or Ebola virus involves
PT delivering a genetic vaccine, which contains a genetic construct encoding
PT antigenic determinants for filovirus, using a particle acceleration
PT device.

Claim 5; Col 25-30; 33pp; English.
XX
XX The patent discloses a method of inducing an immune response to Marburg
CC or Ebola virus glycoprotein which involves delivering a genetic vaccine,
CC containing a genetic construct encoding antigenic determinants for
CC filovirus, using a particle acceleration device. The genetic vaccine for
CC filovirus is created by joining a DNA sequence encoding at least a
CC portion of the filovirus glycoprotein to a promoter effective to promote
CC transcription of the DNA sequence. This method is useful for inducing
CC humoral, cell-mediated and secretory immune responses in the treated
CC individual. The present sequence is a DNA encoding Marburg virus Musoke
CC strain glycoprotein (GP), which is a type I transmembrane protein. This
CC sequence is used in the construction of genetic vaccine against
CC filoviruses

XX

SQ Sequence 2164 BP; 701 A; 536 C; 412 G; 515 T; 0 U; 0 Other;
Query Match 74.0%; Score 1514.8; DB 4; Length 2164;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;
QY 1 ATGAAGACCATATATTTCTGATGTCTCAATTTTAATCCAAAGTATATAAAATCTCCCT 60
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 178
QY 61 GTTTTGAATAATGCTAGTAAACAGCCAACTCAAGATGTAGATTCAAGTGTCTCCGGAAAC 120
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 238
QY 121 CTCCAAAGACAGAGATGTTTCATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT 180
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 298
QY 181 GATTCCCTTTTGAAGCATCTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 358
QY 241 GTTGAGTATACGGAAGGAGAGCCAAACCATGTTTCAATATAAGTGTAAACGATCCC 300
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 418
QY 301 TCTGAAAAATCTTGTCTGATGATGCTCCCGATTAATATCCGATATACCTAAATGTAAA 360
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 478
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGACAGGGGATTCCTCATTTTGG 420
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 538
QY 421 GGGGCATTTTCTGTATGATCGGTTGCTCTTACCAATATGTAACGAGCAAGTCTTC 480
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 598
QY 481 ACTGAAGAAATATAGCAGCTATGATTGTTTAAGAAGAGAGTTCAAGATGATTTTTTCT 540
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 658
QY 541 AGCGAAGGACAGGTTATCGTCATGAACTTGACCTCCCAATTAATATTTGACAGC 600
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 718
QY 601 AGCAATGAAGCGCAGAGAAATGATCGGGATGTTTGGCATCTCTCCAAGATATCAACTCC 660
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 778
QY 661 ACAAACAATCAACATGCCCTCCATCTTTAAACCTCCATCCCTGCCACAGTAACCTCG 720
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 838
QY 721 AGCATTCATCTACAATACTCAAAATTAATCTGCTAAATCTGGAATATGAAACCCAGT 780
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 898
QY 781 AGCGACGATGAGGACCTTTATGATTTTCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 958
QY 841 ACTCTTAATGTAGTCACTGAAACAGAAACAATCGTCAACAATATTTGTCCACTCTTCACTA 900
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1018
QY 901 CATCCAGCACCTCAACAATGAGCAAAAGAGGGCTTTTCATCAACAATGCCACCTCCCTACCA 1060
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1078
QY 961 GAGCACAAATGGAACCGACCCCAACAAACACACAGCAAGCTCTCTCAACAATTAATACA 1020
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1138
QY 1079 GAACTAGACAAAATAACAACTGACAACTGCGTCCATGCCCCCTCATACACTACCA 1138

1021	Qy	ACTCCGACCTATTAACTCTCAAGTACAACCTCAGTACTCTCTTCCCTCCGACCCGGAAC	1081
1139	Db	ATCTCTACTAACCAACACCTCCAAACACAACCTCTCAGCACTCTCTCTGCACCATTTACAAAAC	1198
1081	Qy	ATCACCAATAATGATATACACAACGTGAACTAGCAGAAAGCGAACAAACCAATGCTCAGTTTG	1140
1199	Db	ACACCCANTGACAAACACACAGAGCACAATCACTGAAATGAGGCNAACCATGCGCCCTCG	1258
1141	Qy	AACACAACTCTAGATCCAACAGAAAAATCCCAACACAGGACAGACACCAACAGCACAAAACC	1200
1259	Db	ATAACAACCTCGCTCCCAACGGGAATCCCAACACACAGCAAAAGAGCACACAGACGACAAAAA	1318
1201	Qy	AACATCATCATGAGGACATCAGATATAAACAGCAACACACCCCAACAAATTTCTTCTCCGGAT	1260
1319	Db	GGCCCCGCCACAACGGGCAACCAACACAGCAAAATGACATTTTCAACAGTCTCTCCCCCAACC	1378
1261	Qy	TCTAGTCCGACAAACCCGCCCTCTATATACCTTTTGAAGAAAGAAAGATCGATCTCTCGGAGG	1320
1379	Db	CCCAGCTCGAGTCACACAACATCTTGATATATTCAGAAGAAAGCGNAGTATCCTCTGGAGG	1438
1321	Qy	GAAGGGACATGTTCCCTTTTCTGGATGGGTTTAATAAATGCTCCAAATGATTTTGACCCA	1380
1439	Db	GAAGGGACATGTTCCCTTTTCTGGATGGGTTTAATAAATGCTCCAAATGATTTTGACCCA	1498
1381	Qy	GTTCCTCAATACAAAAACAATCTTTTGATGAATCTCTAGTTTCTGTCGCTCGGCTGAGGAA	1440
1499	Db	GTTCCTCAATACAAAAACAATCTTTTGATGAATCTCTAGTTTCTGTCGCTCGGCTGAGGAA	1558
1441	Qy	GATCAACATGCTCCCCCAATATAGTTTAACTTTATCTTATTTTCTTAATAATAAATGAG	1500
1559	Db	GATCAACATGCTCCCCCAATATAGTTTAACTTTATCTTATTTTCTTAATAATAAATGAG	1618
1501	Qy	AACACTGCCCTACTCTGAGAAATGAGAAATGATGTGATCGAGAGTTTAAAGAAATTCGAGC	1560
1619	Db	AACACTGCCCTACTCTGAGAAATGAGAAATGATGTGATCGAGAGTTTAAAGAAATTCGAGC	1678
1561	Qy	GTTCAGGAGGATGACCTGGCGCAGGCGCTCAGTTGGATACCGTTTTTTGGCCCTCGAANT	1620
1679	Db	GTTCAGGAGGATGACCTGGCGCAGGCGCTCAGTTGGATACCGTTTTTTGGCCCTCGAANT	1738
1621	Qy	GAAGGACTTTTACCTGCTGTTTTTAATTAATAAATCAAAACAATTTGGTCTCGAGTTGAGG	1680
1739	Db	GAAGGACTTTTACCTGCTGTTTTTAATTAATAAATCAAAACAATTTGGTCTCGAGTTGAGG	1798
1681	Qy	CGTCTAGCCAAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA	1740
1799	Db	CGTCTAGCCAAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA	1858
1741	Qy	AGAACATTCTCCTTAAATAGACATGCTATTGACTTTTCTACTCAACAGATGGGAGGA	1800
1859	Db	AGAACATTCTCCTTAAATAGACATGCTATTGACTTTTCTACTCAACAGATGGGAGGA	1918
1801	Qy	ACATGCAAGTGCTTGGACCTGATTTGTTGCATCGGATAGAGACTGTGTCCTCAAAAATATT	1860
1919	Db	ACATGCAAGTGCTTGGACCTGATTTGTTGCATCGGATAGAGACTGTGTCCTCAAAAATATT	1978
1861	Qy	TCAGAGCAAAATTGACCAAAATTAATAAGGACGACAAATAAGAGGAGGACTGTGTTGGGCTCG	1920
1979	Db	TCAGAGCAAAATTGACCAAAATTAATAAGGACGACAAATAAGAGGAGGACTGTGTTGGGCTCG	2038
1921	Qy	GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTTGGGCAATTTGCTACTA	1980
2039	Db	GGTGGTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTTGGGCAATTTGCTACTA	2098
1981	Qy	TTATCCCATAGCTGTCTTGATTGCTCTATCCTGTATTTGTCTGTATCTTTACTAAATATATC	2040
2099	Db	TTATCCCATAGCTGTCTTGATTGCTCTATCCTGTATTTGTCTGTATCTTTACTAAATATATC	2158
2041	Qy	GGATAA 2046	
2159	Db	GGATAA 2164	

RESULT 6	
ACCT71550	
ID ACC71550 standard; DNA; 7778 BP.	
XX AC	
XX ACCT71550;	
XX	
DT 10-JUN-2003 (first entry)	
XX	
XX VRC6701 (pVR1012-Marburg) plasmid.	
XX DE	
XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;	
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;	
KW viral infection; filovirus; circular; cyclic; ds.	
XX OS Synthetic.	
XX	
XX WO2003028632-A2.	
XX PN	
XX 10-APR-2003.	
XX PD	
XX 24-SEP-2002; 2002WO-US030251.	
XX PF	
XX 01-OCT-2001; 2001US-0326476P.	
XX PR	
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.	
XX PA	
XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;	
XX FI	
XX WPI; 2003-371961/35.	
XX DR	
XX New bimodal priming and boosting compositions, useful as viral vaccines,	
XX PT specifically for eliciting an immune response against a filovirus or a	
XX PT disease caused by infection with filovirus.	
XX	
PS Claim 1; Page 185-187; 219pp; English.	
XX	
CC The present invention relates to a bimodal priming composition and	
CC boosting composition for priming and boosting an immune response to an	
CC antigen in an individual. The compositions comprise (a) a priming	
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule	
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza	
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a	
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting	
CC composition comprised of a replication-deficient adenovirus, comprising a	
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus, or	
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or	
CC epitope-bearing domain, or a replication deficient adenovirus selected	
CC from constructs ACC71521-ACC71563. The compositions are useful as viral	
CC vaccines, specifically for eliciting an immune response against a	
CC filovirus or a disease caused by infection with filovirus e.g. Ebola	
CC virus infection. The present sequence has a Marburg glycoprotein (GP)	
CC open reading frame, Musoke strain, cloned into a VRC6700 backbone	
XX	
SQ Sequence 7778 BP; 2148 A; 1892 C; 1719 G; 2019 T; 0 U; 0 Other;	
Query Match 74.0%; Score 1514.8; DB 10; Length 7778;	
Best Local Similarity 83.8%; Pred. No. 0;	
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0	
OY 1 ATGAAGACCATATATTTTCTGATTAGTCTCAATTTTAATCCAAAGTATAAAAACCTCCCT 60	
DB 2034 ATGAAGACCATATTTTCTGATTAGTCTCAATTTTAATCCAAAGTATAAAAACCTCCCT 2095	
OY 61 GTTTTAGAANTGTAGTAAACAGCAACCTCAAGATAGATTAGTGTCGCCGAACC 120	
DB 2094 ATTTTAGAGATAGTACTGTAATAATCAACCCCAAAATGTGGATTGGTATGCTCGCGAAT 2155	
OY 121 CTCCAAAGACAGAGAAGTGTTCATCTGATGGATTTCACATGAGTGGGCAAAAAGTTGCT 180	
DB 2154 CTCAGAAGACAGAGAAGCGTCCATCTGATGGGATTCACATGAGTGGGCAAAAAGTTGCT 2215	
OY 181 GATTCCCCCTTTGGGAAGCATCTAAACGATGGGCTTTTCAGGACAGAGGTGTTCTCCCAGAAC 240	

Db 2214 GATTCCTCCCTTTGGAGGCAATCAAGGATGGGCTTTCAGBACAGGTGTACCTCCCAAGAA 2273
Qy 241 GTTGAGTATACGGAAGAGAGAGCCAAACATGTTACAAATAAGTGTAAACAGACCCCT 300
Db 2274 GTTGAGTACACAGAGGGGAGGAGCCAAACATGCTACANATAAAGTGAACGGATCCC 2333
Qy 301 TCTGGAAAATCCTTGCTGCTGGATCTCCCGAGTAATATCCCGGATATACCTTAATATGTA 360
Db 2334 TCTGGAAAATCCTTGCTGCTGGATCTCCCGAGTAATATCCCGGATATACCTTAATATG 2393
Qy 361 ACTGTTCAATATATCAAGGTCAAAACCTCATGACAGAGGGATGGCCCTCAATTTG 420
Db 2394 ACTATCCATATATCAAGGTCAAAACCTCATGACAGAGGGATGGCCCTCAATTTG 2453
Qy 421 GGGGCAATTTTCTGTATGATCGGGTGGCTCTCAACAATGTATCCCGAGGCAAGGTCTTC 480
Db 2454 GGAGCAATTTTCTGTATGATCGGGTGGCTCTCAACAATGTATCCCGAGGCAAGGTCTTC 2513
Qy 481 ACTGAAGAAATATAGCAGCTATGATTTGAATGAAGCAGTTTCACAGATGATTTTCT 540
Db 2514 ACTGAAGGAAATATAGCAGCTATGATTTGAATGAAGCAGTTTCACAGATGATTTTCT 2573
Qy 541 AGGCAAGGACAGGTTATCGTCAATGAATCTTGACCTCCACCAATATATATGACAGC 600
Db 2574 CGGCAAGGACAGGTTATCGTCAATGAATCTTGACCTCCACCAATATATATGACAGC 2633
Qy 601 AGCAATGAAGCGAGAGAAATGATACGGGATGTTTGGCATCTCTCAAGAAATACAACTCC 660
Db 2634 AGTAACGGAACGCAACGATGACACTGGATGTTTGGGCTCTTCAAGAAATACAACTCT 2693
Qy 661 ACAAAATCAAAATGATCCCTCAATCTCTTAAATCTCCATCCCTGCCACAGTAACCTCG 720
Db 2694 ACAAGGAACCAAAATGATGCTCCGTCCAAAATACCTCCACCACTGCCACAGCCCGTCCG 2753
Qy 721 AGCAATCACTTACAAATCAATTAATTAATCTGTAATCTGTAATCTGTAATCTGTAAT 780
Db 2754 GAGATCAAACTCAACAGCACCCCACTGATGCCCAAACTCAATATACCGAGCCCAAGC 2813
Qy 781 AGCGAGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db 2814 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGGAGAACGAGAACCCACACA 2873
Qy 841 ACTCTTAATGATGATCACTGAAACAGAAACAAATCGTCAACAATATTTGTCACCTCTTCACTA 900
Db 2874 ACTTCTGATCGGGTCACCAAGCAAGGGCTTTTCAACAATGCAAACTCCCTCCACCA 2933
Qy 901 CATCCAGCACTCAACAATGAGCAACAGTAGGATCTCTCCCGACATGCTGTAACT 960
Db 2934 CAACCAAGCACGCCACAGCAAGGAGGAAACCAACCAATTTCCCAAGATGCTGTGACT 2993
Qy 961 GAGCACAATGGAACCGACCCCAACACACACAGCAACAGCTCTCTCAACAATACTAATACA 1020
Db 2994 GAACTAGCAAAATAACAACTGCAACACGGTCCATGCCCCCTCATACACTACCAACA 3053
Qy 1021 ACTCCCACTTAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCCAACCCGCAAC 1080
Db 3054 ATCTCTACTAACAACACCTCCAAACACAACTTCAGCACTCTCTGACCACTTACAAAC 3113
Qy 1081 ATACCAATATATGATACAAACAGTGAACCTAGAGAAAGGAAACCAACCAATGCTCAGTTG 1140
Db 3114 ACCACCAATATGACAAACAGAGCAACATCACTGAAAATGAGCAACCAAGTGCCGCCCTCG 3173
Qy 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACAGGACAGCAACCAACCAACAGCAAC 1200
Db 3174 ATAACACCTGCTCCCAACAGGAAATCCCAACAGCAACAGCAACAGCAACAGCAAC 3233
Qy 1201 AACATCATATGACAGATCATGATATAACAGCAACACCCCAAAATTTCTTCCGGAT 1260
Db 3234 GGGCCCGCACAAACGGCACCAACAGCAACATGAGCATTTTCAACAGTCTCTCCGCCAC 3293
Qy 1261 TCTAGTCCGACAAACCGGCCCTCTTATATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320

Db 3294 CCAGCTCGACTGCACAAACATCTTGTATATTTTCAGAAAGAAAGCAAGTATCTCTGGAGG 3353
Qy 1321 GAAGGCGACATGTTCCCTTTCTGGATGGGTAAATAATGCTCCAAATGATTTTGAACCA 1380
Db 3354 GAAGGCGACATGTTCCCTTTCTGGATGGGTAAATAATGCTCCAAATGATTTTGAACCA 3413
Qy 1381 GTTCCAATATACAAAACAAATCTTTGATGAATCTCTAGTTCTGCTCGCTCGGCTGAGAA 1440
Db 3414 GTTCCAATATACAAAACAAATCTTTGATGAATCTCTAGTTCTGCTCGCTCGGCTGAGAA 3473
Qy 1441 GATCAACATGCTCTCCCAATATATAGTTTAACTTTATCTTATTTTCTTAATAATAATGAG 1500
Db 3474 GATCAACATGCTCTCCCAATATATAGTTTAACTTTTAACTTTTCTTAATAATAATGAG 3533
Qy 1501 AACACTGCCCTCTCTGAGAAATAGAAATGATGTGATGAGAGTTAAAGAAATTTGGAGC 1560
Db 3534 AACACTGCCCTCTCTGAGAAATAGAAATGATGTGATGAGAGTTAAAGAAATTTGGAGC 3593
Qy 1561 GTTCAGGAGGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTGTGGCCCTGGAAAT 1620
Db 3594 GTTCAGGAGGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTGTGGCCCTGGAAAT 3653
Qy 1621 GAAGGACTTTTACACTGCTGTTTTTAAATAATAATAATAATAATAATAATAATAATA 1680
Db 3654 GAAGGACTTTTACACTGCTGTTTTTAAATAATAATAATAATAATAATAATAATAATA 3713
Qy 1681 CGTCTAGCCAAATCAAACTGCCAAATCTTTTGAATCTTTTGAAGAGTCACAACTGAGAA 1740
Db 3714 CGTCTAGCCAAATCAAACTGCCAAATCTTTTGAATCTTTTGAAGAGTCACAACTGAGAA 3773
Qy 1741 AGAATCTCTCTTAATCAATAGACATGCTATTTGACCTTTTCTACTCAAGATGGGAGGA 1800
Db 3774 AGAATCTCTCTTAATCAATAGACATGCTATTTGACCTTTTCTACTCAAGATGGGAGGA 3833
Qy 1801 ACATGCAAAAGTCTTGGACCTGATTTGCAATCGGATAGAGACTTGTCCAAAATAAT 1860
Db 3834 ACATGCAAAAGTCTTGGACCTGATTTGCAATCGGATAGAGACTTGTCCAAAATAAT 3893
Qy 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAGAGGGGACTGTTGGGGCTG 1920
Db 3894 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAGAGGGGACTGTTGGGGCTG 3953
Qy 1921 GGTGTAATAGTGGAGACATCCGACTGGGTGTTCTTACTAACTTGGGCAATTTGCTACTA 1980
Db 3954 GGTGTAATAGTGGAGACATCCGACTGGGTGTTCTTACTAACTTGGGCAATTTGCTACTA 4013
Qy 1981 TTATCCATAGCTGCTTGGATGCTATCTGATTTGCTGATTTGCTGATTTTACTAAATATATC 2040
Db 4014 TTATCCATAGCTGCTTGGATGCTATCTGATTTGCTGATTTTACTAAATATATATC 4073
Qy 2041 GGATAA 2046
Db 4074 GGATAA 4079
RESULT 7
AAZ87211
ID AAZ87211 standard; cDNA; 11460 BP.
XX AAZ87211;
AC AAZ87211;
XX 15-SEP-2003 (revised)
DT 08-MAY-2000 (first entry)
XX VEE replicon comprising Marburg virus genes.
DE VEE virus replicon; Venezuelan equine encephalitis; Marburg virus; MBGV;
KW filovirus; glycoprotein; GP; nucleoprotein; NP; structural protein; VP40;
KW VP30; VP30; VP24; genetic vaccine; antigen delivery;
KW Marburg haemorrhagic fever; cyclic; circular; ss.
XX Marburg virus; str. Musoke.
OS Venezuelan equine encephalitis virus.

OS Chimeric.

XX Key Location/Qualifiers

FT CDS 104..2182

FT /tag= a

FT /product= "Marburg virus nucleoprotein (NP, AAY77128)"

FT 2944..3933

FT /tag= b

FT /product= "Marburg virus structural protein VP35 (AAY77130)"

FT 4567..5478

FT /tag= c

FT /product= "Marburg virus structural protein VP40 (AAY77129)"

FT 5940..7985

FT /tag= d

FT /product= "Marburg virus glycoprotein (GP, AAY77127)"

FT 8864..9697

FT /tag= e

FT /product= "Marburg virus structural protein VP30 (AAY77131)"

FT 10200..10961

FT /tag= f

FT /product= "Marburg virus structural protein VP24 (AAY77132)"

XX WO200000616-A2.

XX 06-JAN-2000.

XX 21-JUN-1999; 99WO-US014174.

XX 29-JUN-1998; 98US-0091403P.

XX (USNE-) US MEDICAL RES INST INFECTIOUS DISEASES.

XX Hevey MC, Negley DL, Pushko P, Smith JF, Schmaljohn AL;

XX WPI; 2000-160676/14.

XX P-PSDB; AAY77127, AAY77128, AAY77129, AAY77130, AAY77131, AAY77132, AAY77133.

XX Novel Marburg virus vaccines used to induce an immune response against the infection in nonhuman primates.

XX Claim 1; Page 56-57; 57pp; English.

XX The invention relates to novel Marburg virus (MBGV) vaccines. The vaccine of the invention comprises a Venezuelan equine encephalitis virus (VEE) replicon containing copies of the genes encoding Marburg virus (Musoke strain) glycoprotein (GP), a deletion mutant GP (GP-delta-TM), nucleoprotein (NP), and structural proteins VP40, VP35, VP30 and VP24. The Marburg virus genes replace the VEE virus structural protein genes; the result is a self-replicating RNA molecule that encodes its own replicase and transcriptase functions, and in additional makes abundant quantities of the Marburg virus proteins. When replicon RNA is transfected into eukaryotic cells, along with two helper RNAs that express the VEE virus structural proteins, the replicon RNA is packaged into VEE virus-like particles by the VEE virus structural proteins, which are provided in trans. Since the helper RNAs lack packaging signals necessary for further propagation, the resulting VEE replicon particles (VRPs) which are produced are infectious for one cycle but are defective thereafter. On VNP infection of a cell, an abortive infection occurs whereby the cell produces the Marburg virus proteins, is ultimately killed by the infection, but does not produce any viral progeny. The VEE replicon provides a potent tool for vaccination with Marburg virus antigens. The replicons, vectors and constructs are used to produce vaccines against Marburg virus (MBGV) infection (Marburg haemorrhagic fever) in mammals, to elicit immune responses against Marburg antigens, to confer protective immunity, and to reduce disease symptoms and reduce the severity of disease. Studies of non-human primates vaccinated with the replicon indicate that the vaccine will be efficient in protecting humans against Marburg virus. Prior art Marburg virus vaccination

CC strategies have used formalin-inactivated Marburg virus, which are only partially successful at protecting against Marburg virus infection. The CC present sequence represents a cDNA corresponding to a VEE virus replicon encoding Marburg virus GP, GP-delta-TM, NP, VP40, VP35, VP30 and VP24. CC (Updated on 15-SEP-2003 to standardise OS field)

XX Sequence 11460 BP; 3674 A; 2382 C; 2145 G; 3259 T; 0 U; 0 Other;

SQ Query Match 74.0%; Score 1514.8; DB 3; Length 11460; Best Local Similarity 83.8%; Pred. No. 0; Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

Qy 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTATTCCTCAAGTATATAAACTCTCCCT 60

Db 5940 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTAATTCAGGAGCAAAAATCTCCCC 5999

Qy 61 GTTTTGAAGATTCCTAGTAAACAGCAACCTCAAGATGTAGATTCAGTGTGCTCCGGAACC 120

Db 6000 ATTTTGAAGATAGCTAGTAATATCAACCCCAAAATGCGTATGCTCCGGAACCT 6059

Qy 121 CTCCTCAAGAGCAGAAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 180

Db 6060 CTCCTCAAGAGCAGAAAGATGTTTCATCTGATGGGATTTACACTGAGTGGGCAAAAAGTTGCT 6119

Qy 181 GATTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240

Db 6120 GATTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 6179

Qy 241 GTTGAGTATACGGAGGAGGAGAGCCAAACATGTTACAATATAAGTGTACAGACCCCT 300

Db 6180 GTTGAGTATACGGAGGAGGAGAGCCAAACATGTTACAATATAAGTGTACAGACCCCT 6239

Qy 301 TCTGAAAAATCTTTGCTGCTGGATCTCTCCAGTAAATATCCGCGATTACCTTAAATGTAAA 360

Db 6240 TCTGAAAAATCTTTGCTGCTGGATCTCTCTACCAACATCCGCTGACTATCTTAATGTCAA 6299

Qy 361 ACTGTTTATCATATATCAAGGTCAAAACCTCATATGCACAGGGGATGTCCTCATTTTGGG 420

Db 6300 ACTATCCATCATATTTCAAGGTCAAAACCTCATATGCACAGGGGATGTCCTCATTTATGG 6359

Qy 421 GGGGCATTTTCTGTATGATCGGTTGCTCTCAACAATGTACCGAGGCAAGGTCTTC 480

Db 6360 GGAGCATTTTCTGTATGATCGGTTGCTCTCAACAATGTACCGAGGCAAGGTCTTC 6419

Qy 481 ACTGAAGGAAATATAGCAGCTATGATTGTTTAAATAGACAGTTTCACAGAAATGATTTTTCT 540

Db 6420 ACTGAAGGAAATATAGCAGCTATGATTGTTTAAATAGACAGTTTCACAGAAATGATTTTTCT 6479

Qy 541 AGGCAAGGACAGGTTATCGTCACTGAACTTGACCTCCACCAATTAATATTTGGCAAGC 600

Db 6480 CGGCAAGGACAGGTTATCGTCACTGAACTTGACCTCCACCAATTAATATTTGGCAAGC 6539

Qy 601 AGCAATGAAAGCGCAGAGAAATGATACGGGATGTTTGGCATCTCTTAAACCTCCATCCCTGCCACAGTAACTCC 660

Db 6540 AGTAAAGGAAAGCGCAGAGAAATGATACGGGATGTTTGGCATCTCTTAAACCTCCATCCCTGCCACAGTAACTCC 6599

Qy 661 ACAACCAATCAAAATGCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACTCC 720

Db 6600 ACAAGAACCAAAATGCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACTCC 6659

Qy 721 AGCAATCACTCTACAAATGCTCAAAATTAATGCTGCTAAATCTGGAAGTATGAAACCAAGT 780

Db 6660 GAGATCAAACTCAAGCAGCAGCCCAACTGATGTCACCAAACTCAATACCAAGGACCCCAAGC 6719

Qy 781 AGCCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840

Db 6720 AGTATGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 6779

Qy 841 ACTCTTAATGATGAGTCACTGAAACAGAAACAATCGTCAACAATTTGTCCACTCTTCTCACTA 900

Db 6780 ACTTCTGATGGGTCAACCAAGAGGGCTTTTCATCAACAATGCCACCCACTCTTCTCACTA 6839

Qy 901 CATCCAGCACCTCACAAACATGAGCAAAAACAGTAGGAATCTCTCCCGACATGCTGTAAC 960

Db 6840 CAACAGCAGCCACGACGAGGAGAAACAAACAAACCAATCCCAAGATGCTGTGACT 6899
Qy 961 GAGCACAATGGAACCGACCCCAACCAACCAACGACGCTCTCTCAACAATACTAATACCA 1020
Db 6900 GAACTAGACAAATAAACAACACTGCAACACCGTCCATGCCCCCTCATAACTACCTACCACA 6959
Qy 1021 ACTCCACCTATACACTCTCAAGTACACCTCTAGTACTCTTCCCTCCGACCCGCAAC 1080
Db 6960 ATCTCTACTAAACACCTCCAAACACAACTTTCAGCACTCTCTCTGCACCAATTAACAAC 7019
Qy 1081 ATACCAATAATATACACACGCTGAATAGCAGAAAGCGAAACAAACCAATGCTCAGTTG 1140
Db 7020 ACCACCAATGACACACACAGACACATCACTGAAATGAGCAACCAAGTCCCTCTCG 7079
Qy 1141 AACCAACTCTAGATCCAAAGAAAATCCCAACAGGACCAAGACACCAACGACCAAC 1200
Db 7080 ATAACAACCTGCTCCAAAGGAAATCCCAACAGCAGCAAGAGACACGACGACAAA 7139
Qy 1201 AACATCATGACGACATCAGATATACAGAGCAAAACACCCCAACAATTTCTTCCGGAT 1260
Db 7140 GGGCCCGCCACACAGGACCAACACGACAAATGAGCAATTCACGACGCTCTCCGCCACC 7199
Qy 1261 TCTAGTCGACACCGCCCTCTATATACCTTTAGAAAGAAACGATCGATCCTCTGGAGG 1320
Db 7200 CCCAGCTGCACTGCACAACTCTTGATATATTTTCAGAAAGCGAAGTATCTCTGGAGG 7259
Qy 1321 GAAGCGCAGATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAAATTTGATTTGACCCA 1380
Db 7260 GAAGCGCAGATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAAATTTGATTTGACCCA 7319
Qy 1381 GTTCCAAATACAAAACAACTCTTGTAGAAATCCTCTAGTTCTGGTCTCGGCTGAGGAA 1440
Db 7320 GTTCCAAATACAAAACAACTCTTGTAGAAATCCTCTAGTTCTGGTCTCGGCTGAGGAA 7379
Qy 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTCTTATTTCTTAAATAAGTAG 1500
Db 7380 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTCTTATTTCTTAAATAAGTAG 7439
Qy 1501 AACACTGCTACTCTGGAGAAAATCAGAAATGATTTGTATGATGAGATTTAAAGATTTGGAGC 1560
Db 7440 AACACTGCTACTCTGGAGAAAATCAGAAATGATTTGTATGATGAGATTTAAAGATTTGGAGC 7499
Qy 1561 GTTCAGGAGGATGACTGGCCGCGAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAATT 1620
Db 7500 GTTCAGGAGGATGACTGGCCGCGAGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAATT 7559
Qy 1621 GAAGGACTTTACATGCTGTTTTTAATTAATAATCAAAACAAATTTGGTCTGAGGTTGAGG 1680
Db 7560 GAAGGACTTTACATGCTGTTTTTAATTAATAATCAAAACAAATTTGGTCTGAGGTTGAGG 7619
Qy 1681 CGTCTAGCAATCAAACTGCCAAATCCTTTGGAATCTTTATTTAGAGTCCAACTGAGGAA 1740
Db 7620 CGTCTAGCAATCAAACTGCCAAATCCTTTGGAATCTTTATTTAGAGTCCAACTGAGGAA 7679
Qy 1741 AGAATCTCTCTTAATCAATAGACATGCTATTTGATTTCTTACTCACAAGATGGGAGGA 1800
Db 7680 AGAATCTCTCTTAATCAATAGACATGCTATTTGATTTCTTACTCACAAGATGGGAGGA 7739
Qy 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGATAGAGACTTGTCCAAAATATT 1860
Db 7740 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGATAGAGACTTGTCCAAAATATT 7799
Qy 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACCAACAAAAGAGGGGACTGTGTGGGCTCG 1920
Db 7800 TCAGAGCAAAATGACCAAAATTAATAAGGACCAACAAAAGAGGGGACTGTGTGGGCTCG 7859
Qy 1921 GGTGTAATGTTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTGCTACTA 1980
Db 7860 GGTGTAATGTTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTGCTACTA 7919
Qy 1981 TTATCCATAGCTTGTGATGCTCTATCCCTGATTTGTGCTATCTTTACTAAATATATC 2040

Db 7920 TTATCCATAGCTGCTGTGATTTGCTCTATCTCTGATTTGTGCTATCTTTACTAAATATC 7979
Qy 2041 GGATAA 2046
Db 7980 GGATAA 7985
RESULT 8.
AAZ51039
ID AAZ51039 standard; DNA; 2100 BP.
XX AAZ51039;
AC AAZ51039;
XX 05-JUN-2000 (first entry)
XX Marburg virus envelope glycoprotein DNA.
XX Marburg virus envelope glycoprotein; viral glycoprotein;
XX pseudotyped retrovirus; MLV; Moloney murine leukaemia virus; antibody;
XX transduction; screening agent; immunological agent;
XX pharmacological agent; ss.
XX Marburg virus.
XX Key Location/Qualifiers
FT CDS 10..2052 /*tag= a
FT /product= "Marburg virus envelope glycoprotein"
XX WO200008131-A2.
XX 17-FEB-2000.
XX 04-AUG-1999; 99WO-US017702.
XX 04-AUG-1998; 98US-0095242P.
XX 15-DEC-1998; 98US-0112405P.
XX (PURD) PURDUE RES FOUND.
XX Sanders DA, Kuhn RJ, Jeffers SA, Sharkey CM, North CL;
XX Fischbach MA;
XX WPI; 2000-224030/19.
XX P-PSDB; AAY70075.
XX Cells that produce inventive pseudotyped retroviruses having a broad host
XX range useful for introducing nucleotide sequences into target cells.
XX Example 11; Page 60-61; 65pp; English.
XX The present sequence encodes the Marburg virus envelope glycoprotein.
XX Plasmid pMBGP1 was produced by cloning into the plasmid pSP72 nucleotide
XX sequences corresponding to nucleotides 5931-8033 from the Marburg virus
XX genome. This is used for generation of stable cell lines transiently
XX producing Marburg-MLV (Moloney murine leukaemia virus) pseudotyped
XX retroviruses with a broad host range. These cells having different viral
XX glycoproteins in its lipid bilayer with a pseudotyped retrovirus is
XX transduced with a desired ribonucleotide sequence. This is used to
XX identify screening agents effective in blocking viral entry into a cell.
XX These agents may be immunological agents like monoclonal or polyclonal
XX antibodies. The pharmacological agents include proteins, peptides or
XX various chemical agents. The pseudotyped retrovirus may be useful in
XX methods of identifying cell surface receptors that allow viral entry
XX Sequence 2100 BP; 670 A; 532 C; 408 G; 490 T; 0 U; 0 Other;
SQ
Query Match 73.3%; Score 1498.8; DB 3; Length 2100;
Best Local Similarity 83.6%; Pred. No. 0;
Matches 1711; Conservative 0; Mismatches 332; Indels 3; Gaps 1;
Qy 1 ATGAAGACCATATATTTTCTGATTTAGTCTCATTTTATCCAAAGTATATAAACTTCCTCCT 60

Dbb 10 ATGAAGACCA CATGTTTCTTATCAAGTCTTATCTTAAATTCAAGGGACAAAAATCTCCCC 69
Qy 61 GTTTTATAGAAATGCTAGTAAACAGCCCAAGCTCAAGATGTAGATTTCAGTGTGCTCGGGAACC 120
Dbb 70 ATTTTAGAGATAGCTAGTAAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCGGGAAC 129
Qy 121 CTCCTCAAGACAGAAAGATGTTTCAATCTGATGGGATTTTACATGAGTGGGCAAAAATGTTGCT 180
Dbb 130 CTCAGAGACAGAGAAGCGTCCATCTGATGGGATTTCAACATGAGTGGGCAAAAATGTTGCT 189
Qy 181 GATTCCCTTTTGGAGCATCTAAGCATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Dbb 190 GATTCCCTTTTGGAGCATCAAGGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 249
Qy 241 GTTGTAGTACGGAGGAGAGAGCCAAACATGTTACAATAAAGTGTAAACAGACCC 300
Dbb 250 GTTGTAGTACAGAGGGGAGAGCCAAACATGTTACAATAAAGTGTAAACAGATCCC 309
Qy 301 TCTGGAAAATCTTCTGCTGTGATCCTCCCAAGTAATATCCCGGATTTACCTTAAATGTAAA 360
Dbb 310 TCTGGAAAATCTTCTGCTGTGATCCTCCCAAGTATCCCGGATTTACCTTAAATGTAAA 369
Qy 361 ACTGTTTCATATATTCAGGTCAAAACCTCATGACAGGGGATTTGCCCTCCATTTGTTGG 420
Dbb 370 ACTATCCATCATATTCAGAGGTCAAAACCTCATGACAGGGGATTTGCCCTCCATTTATGG 429
Qy 421 GGGGCAATTTTCTGTATGATCGGCTTGCCTCTACAACAATGTATCCAGGCAAGTCTTTC 480
Dbb 430 GGAGCATTTTCTGTATGATCGCATTCGCTTCCACCAATGTATCCAGGCAAGTCTTTC 489
Qy 481 ACTGAAGGAATATATAGCAGCTATGATTTTAAATAGACAGTTCACAGAAATGATTTTCT 540
Dbb 490 ---GAAGGGAACATAGCAGCTATGATTTTCAATAGACAGTTCACAGAAATGATTTTCTCG 546
Qy 541 AGGCAAGGACAGGTATTCGTACATGAACTTGACCTCCACCAATAAATATTTGGACAGC 600
Dbb 547 CGGCAAGGACAGGTATTCGTACATGAACTTGACCTCCACCAATAAATATTTGGACAA 606
Qy 601 AGCAATGAACCCAGAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAACCTCC 660
Dbb 607 AGTAACGGAACGCAACCAATGACACTGGATGTTTGGGGCTTTCAGAAATACAAATCT 666
Qy 661 ACAAACATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCCAAGTAACTCCG 720
Dbb 667 ACAAGAAACCAACATGTGCTCCGTCCCAAAATACCTCCACCACTGCCCCACAGCCGTCG 726
Qy 721 AGCATTCATCTACAAATACCAAAATTAATCTGCTAAATCTGGAACTATGACCCCAAGT 780
Dbb 727 GAGATCAAACTCACAAAGCACCCCAACTGATGCCCAAACTCAATACCAAGACCCCAAGC 786
Qy 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Dbb 787 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACAGAGACCCACACA 846
Qy 841 ACTCTTAATGATGATCTGAAACAGAAAACAAATGTCACAAATATGTCACATCTCTTCACTA 900
Dbb 847 ACTTCTGATGCGGTCAACCAAGCAAGGGCTTTCATCAACAATGCCACCACTCCCTCAACA 906
Qy 901 CATCAAGCACTCACAAATGAGCAAAACAGATAGCAATCTTCCCGACATGCTGTAACT 960
Dbb 907 CAACCAAGCAAGGAGGAGGAAACCAACAAACCAATTTCCCAAGATGCTGTGACT 966
Qy 961 GAGCAATGGAACCGACCCCAACACACAACAGCAAGCTCTCAACAATACTAATACA 1020
Dbb 967 GAACTGACAAATAATACAACTGCAACCGTCCATGCCCCCTCATTAACACTACCAACA 1026
Qy 1021 ACTCCCACTATAACATCTCTCAAGTACAACTCTAGTACTCTTCCCTCCCAACCGGCAAC 1080
Dbb 1027 ATCTCTACTAACAACACCTCCAAACACAACTTTCAGCACTCTCTCTGACCACTTACAAAAC 1086
Qy 1081 ATCACCATAATGATACACAGGTGAATAGCAGAAAGGCAACCAACCAATGCTCAGTTG 1140
Dbb 1087 ACCACCAATGACAAACACACAGAGCAATCACTGAAATATGACAAACAGTGGCCCCCTCG 1146

Qy 1141 AACACAACTCTAGATCCAAACAGAAAAATCCCAACCAAGGACAAAGACACCAACAGCACAAACC 1200
Dbb 1147 ATAACAACCTCTCCATCCAAACGGGAAATCCCAACCAAGGACAAAGACACCAAGCAAAAAA 1206
Qy 1201 AACATCATCATGACGACATCAGATATAACAGCAAAACACCCCAACAAATTTCTTCCCGAT 1260
Dbb 1207 GGGCCCCCAACACGGACCAACACACGACAAATGAGCATTTTCAACGCTCTCCCCCACC 1266
Qy 1261 TCTAGTCCGACAAACCCCTCTCTATATATCTTTAGAAAGAAACGATCGATCCTCTGGAGG 1320
Dbb 1267 CCAAGCTCGATGACACAACTCTTGATATTTTCAAGAAAGCGAGTATCTCTGGAGG 1326
Qy 1321 GAAGGCGACATGTTTCTCTGATGGGTAAATAAATGCTCCAAATGATTTTGAACCA 1380
Dbb 1327 GAAGGCGACATGTTTCTCTGATGGGTAAATAAATGCTCCAAATGATTTTGAACCA 1386
Qy 1381 GTTCCAAATACAAAACAATCTTTGATGAATCCTCTAGTTCTGCTGCTCGGCTGAGGAA 1440
Dbb 1387 GTTCCAAATACAAAACAATCTTTGATGAATCCTCTAGTTCTGCTGCTCGGCTGAGGAA 1446
Qy 1441 GATCAACATGCTCCCAATATTTAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG 1500
Dbb 1447 GATCAACATGCTCCCAATATTTAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG 1506
Qy 1501 AACACTGCTACTCTCGAGAAAATGAGATGATTTGATGACAGATTTAAGAAATTTGGAGC 1560
Dbb 1507 AACACTGCTACTCTCGAGAAAATGAGATGATTTGATGACAGATTTAAGAAATTTGGAGC 1566
Qy 1561 GTTCAAGAGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTGGCTGCTGCTGGAAT 1620
Dbb 1567 GTTCAAGAGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTGGCTGCTGGAAT 1626
Qy 1621 GAAGCACTTTACACTGCTGTTTAAATTAATAAATCAAAACAATTTGGCTCGCAGGTTGAGG 1680
Dbb 1627 GAAGCACTTTACACTGCTGTTTAAATTAATAAATCAAAACAATTTGGCTCGCAGGTTGAGG 1686
Qy 1681 CGTCTAGCCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 1740
Dbb 1687 CGTCTAGCCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA 1746
Qy 1741 AGAACATTTCTTTAAATAGACATGCTATTGACCTTCTACTCACAGATGGGAGGA 1800
Dbb 1747 AGAACATTTCTTTAAATAGACATGCTATTGACCTTCTACTCACAGATGGGAGGA 1806
Qy 1801 ACATGCAAAAGTCTTGGACCTGATTTGATCGGATAGAAAGTCTTGTCCAAAAATAT 1860
Dbb 1807 ACATGCAAAAGTCTTGGACCTGATTTGATCGGATAGAAAGTCTTGTCCAAAAATAT 1866
Qy 1861 TCAGACAAATTTGACCAAAATTAATAAAGGACGAAACAAAAAGAGGGAAGTGGTTGGGCTG 1920
Dbb 1867 TCAGACAAATTTGACCAAAATTAATAAAGGACGAAACAAAAAGAGGGAAGTGGTTGGGCTG 1926
Qy 1921 GGTGTTAAATGGTGGACATCCGACTGGGTGTTCTTAACTTAACTTGGGCAATTTGCTACTA 1980
Dbb 1927 GGTGTTAAATGGTGGACATCCGACTGGGTGTTCTTAACTTAACTTGGGCAATTTGCTACTA 1986
Qy 1981 TTATCCATAGTCTGTTGATTTGCTATCTCTGATTTGCTGATTTGCTTAACTTAACTATATC 2040
Dbb 1987 TTATCCATAGTCTGTTGATTTGCTATCTCTGATTTGCTGATTTGCTTAACTTAACTATATC 2046
Qy 2041 GGATAA 2046
Dbb 2047 GGATAA 2052

RESULT 9
ABT13454
ID ABT13454 standard; DNA; 2051 BP.
XX ABT13454;
AC ABT13454;
XX 30-JAN-2003 (first entry)
DT

XX DNA encoding a chimeric filovirus protein MBGV-GP1/GP2.
XX
XX Virucide; vaccine; chimeric filovirus glycoprotein; GP; GP1; GP2;
KW immune response; Ebola; Marburg virus; vaccine; gene; ds.
XX
XX Unidentified.
OS Chimeric.
XX
XX WO200279239-A2.
FN
XX
XX 10-OCT-2002.
PD
XX
XX 31-JAN-2002; 2002WO-US003339.
PF
XX
XX 31-JAN-2001; 2001US-0267522P.
PR
XX
XX (USSA) US ARMY MEDICAL RES INST INFECTIOUS DISE.
PA
XX
XX Grogan CC, Hevey MC, Schmaljohn AL;
PI
XX
XX WPI; 2003-040651/03.
DR
XX
XX P-PSDB; ABJ18476.
DR
XX
XX New chimeric filovirus glycoprotein (GP) protein comprising GP1 and GP2,
PT useful for inducing an immune response against infection of different
PT filoviruses, specifically against both Ebola and Marburg viruses.
PT
XX
XX Disclosure; Page 81-84; 94pp; English.
PS
XX
XX The invention relates to a chimeric filovirus glycoprotein (GP) protein
CC comprising GP1 and GP2, where GP1 is from a filovirus different than that
CC of GP2. The chimeric filovirus GP protein is useful for inducing an
CC immune response against infection of different filoviruses, specifically
CC against both Ebola and Marburg viruses by being used as a vaccine. This
CC polynucleotide sequence represents a DNA encoding a chimeric filovirus
CC protein of the invention
CC
XX
XX Sequence 2051 BP; 655 A; 520 C; 400 G; 476 T; 0 U; 0 Other;
SQ

Query Match 72.6%; Score 1485.6; DB 10; Length 2051;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1705; Conservative 0; Mismatches 329; Indels 12; Gaps 1;

Qy 1 ATGAAGACCATATTTTCTGATGATGCTCAATTTTAAATCCAAAGTATATAAAATCTCCCT 60
Db 10 ATGAAGACCATATTTTCTGATGATGCTCAATTTTAAATCCAAAGTATATAAAATCTCCCT 69

Qy 61 GTTTTAGAATGCTAGTAACAGCAACCTCAAGATGATGATGATGATGATGATGATGATGATGAT 120
Db 70 ATTTTAGAGATAGCTAGTAATTAATCAACCCCAAAATGTGGATGCTCGGAACT 129

Qy 121 CTCCAAAGACAGAAGATGTTTCACTGATGGATTTTCACTGATGATGATGATGATGATGATGATGAT 180
Db 130 CTCGAAGACAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 189

Qy 181 GATTCCCTTTTGGAGCATCTAAACGATGGCTTTTCAGGACAGATGTTTCTCCCAAGAAC 240
Db 190 GATTCCCTTTTGGAGCATCTAAACGATGGCTTTTCAGGACAGATGATGATGATGATGATGATGAT 249

Qy 241 GTTGAT 300
Db 250 GTTGAT 309

Qy 301 TCTGAAATCTTGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 310 TCTGAAATCTTGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 369

Qy 361 ACTGTTTCATATATTCAGGTCATAACCTCATGACAGGGGATGATGATGATGATGATGATGATGAT 420
Db 370 ACTATCCATCATATTCAGGTCATAACCTCATGACAGGGGATGATGATGATGATGATGATGATGAT 429

Qy 421 GGGGCAATTTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480

Db 430 GGAGCATTTTCTGTAATGATCGCAATTT-----ATGTACCGAGGCAAGTCTTTC 477
Qy 481 ACTGAAGGAATATPAGCAGCTATGATTGTTTAATAAGACAGTTTCAAGAAATGATTTTCT 540
Db 478 ACTGAAGGAACATPAGCAGCTATGATTGTTCAATAAGACAGTGCACAAAATGATTTTCTCG 537
Qy 541 AGCAAGGACAGAGTTTATGTCATCATGAACCTTGACCTCCACCAATAAATATTGACAGAC 600
Db 538 CGGCAAGGACAGAGGTTACCGTCAATATGAATCTGATTTCTACTAATAAATATTGACAGAT 597
Qy 601 AGCAATGAACGACAGAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAACTCC 660
Db 598 AGTAACGGAACGCAACGAATGACATCTGGATGTTTCGGCGCTCTTCAAGAAATACAAATCT 657
Qy 661 ACAAAACAATCAAAACATGCCCTCCATCTCTTTAAACCTCCATCCCTGCCACAGTAATCTCG 720
Db 658 ACAAGGAACCAACATGCTCGTCCAAATACCTCCACCACTGCCACAGCCGCTCG 717
Qy 721 AGCATCTCTACAAATACCTCAAAATTAATCTGCTAAATCTGGAATCTGAACCCCAAGT 780
Db 718 GAGATCAAACTCACAAGCACCCCACTGATGTCACCAAACTCAATACCCAGACCCCAAGC 777
Qy 781 AGGACCATGAGGACCTTATGATTTTCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db 778 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 837
Qy 841 ACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAATAATTTGTCCACTCTCTTCACTA 900
Db 838 ACTTCTGATGGGTCCACCAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCACC 897
Qy 901 CATCAAGCACTCTCAACAATGAGCAAAACAGTAGGAATCTTCCGACATGCTGTAACT 960
Db 898 CAACCAAGCAGCCACAGCAAGGAGGAAACAACAACCAATTTCCCAAGATGCTGTGACT 957
Qy 961 GAGCACAAATGGAACCGACCCCAACAACAACCAAGCAAGCTCTCAACAATACTAATAACA 1020
Db 958 GAACTAGCAAAAATAACAACATGCAACACCGTCCATGCCCCCTCATACACTACCAACA 1017
Qy 1021 ACTCCCACTTATAACACTCTCAAGTACAACCTCAGTACTCTCTCCCTCCCAACCCGCAAC 1080
Db 1018 ATCTCTACTAACAACACTCCAAACAACAATTCAGCACTCTCTCTGCACTTACCAAAAC 1077
Qy 1081 ATCAACCAATATGATACACAACGTAAGTACAGAGAAAGCGAAACAACAATGCTCAGTTG 1140
Db 1078 ACCACCAATGACACACACAGAGCACAATCACTGAAAATGAGCAAAACGATGCCCTCG 1137
Qy 1141 AACCAACTCTAGATCCAAAGAAAATCCCAACAGCAAGCAAGCAACCAAGCAGACACACC 1200
Db 1138 ATAAACAACCTCTGCTCCAAAGGAAATCCCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1197
Qy 1201 AACATCATATGACGACATCAGATATATAACAGCAAAACACCCCAACAATAATCTTCTCCG 1260
Db 1198 GGCCCCGCCACAGCGCACCAACAACAGCAAAATGAGCATTTTCAACAGTCTCTCCCCAC 1257
Qy 1261 TCTAGTCCGACAAACCCGCTCTCTATATACTTTAGAAAGAAACGATCGATCTCTGAGG 1320
Db 1258 CCCAGCTCGATGCAACAACATCTTGTATATTTCAAGAAAGAGCGATCGATCTCTGAGG 1317
Qy 1321 GAAGGCGACATGTTCCCTTTCTGGATGGGTAAATAATGCTCCAAATGATTTTGACCA 1380
Db 1318 GAAGGCGACATGTTCCCTTTCTGGATGGGTAAATAATGCTCCAAATGATTTTGACCA 1377
Qy 1381 GTTCCAAATACAAAACAATCTTTGATGAATCTCTAGTCTGGTGCCTCGGCTGAGGAA 1440
Db 1378 GTTCCAAATACAAAACAATCTTTGATGAATCTCTAGTCTGGTGCCTCGGCTGAGGAA 1437
Qy 1441 GATCAACATGCTCCCCCAATATTAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG 1500
Db 1438 GATCAACATGCTCCCCCAATATTAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG 1497
Qy 1501 AACCTGCTCTCTGAGGAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560

D	b	1498	AACACTGCCTACTCTGGAGAAAATGAGAAATGATTTGTGATGACAGAGTTAAGAAATTTGGAGC	1557
Q	y	1561	GTTTCAGGAGGATGACCTCGCGCGAGGCTCAGTTGGATACCGTTTTTTGGCCCTCGGAAT	1620
D	b	1558	GTTTCAGGAGGATGACCTCGCGCGAGGCTCAGTTGGATACCGTTTTTTGGCCCTCGGAAT	1617
Q	y	1621	GAAGGACCTTTACACTGCTGTTTAAATTAATAAATCAAAACAATTTGGTCTGCAAGTTGAGG	1680
D	b	1618	GAAGGACCTTTACACTGCTGTTTAAATTAATAAATCAAAACAATTTGGTCTGCAAGTTGAGG	1677
Q	y	1681	CGTCTAGCAATCAAACTGCGCAATCCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA	1740
D	b	1678	CGTCTAGCAATCAAACTGCGCAATCCCTTGGAACTCTTATTTGAGAGTCACAACTGAGGAA	1737
Q	y	1741	AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTTCTACTCACAAAGATGGGAGGA	1800
D	b	1738	AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTTCTACTCACAAAGATGGGAGGA	1797
Q	y	1801	ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAACTTTGTCCAAAAATATT	1860
D	b	1798	ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAACTTTGTCCAAAAATATT	1857
Q	y	1861	TCAGAGCAAAATGACCAATTAATAAAGACCAACAAAAGAGGGGACTGTTGGGCTCG	1920
D	b	1858	TCAGAGCAAAATGACCAATTAATAAAGACCAACAAAAGAGGGGACTGTTGGGCTCG	1917
Q	y	1921	GTTGTAAATGTTGACATCCGACTCGGGGTGTTCTTACTAACTTGGGCAATTTTGCTACTA	1980
D	b	1918	GTTGTAAATGTTGACATCCGACTCGGGGTGTTCTTACTAACTTGGGCAATTTTGCTACTA	1977
Q	y	1981	TTATCCATAGCTGTTGATTTGCTCTATCCCTGATTTTGTCTATCTTTTACTTAATATATC	2040
D	b	1978	TTATCCATAGCTGTTGATTTGCTCTATCCCTGATTTTGTCTATCTTTTACTTAATATATC	2037
Q	y	2041	GGATAA 2046	
D	b	2038	GGATAA 2043	
RESULT 10				
ID	ADM48345	standard; DNA; 19112 BP.		
AC	ADM48345;			
DT	18-NOV-2004	(first entry)		
DE	Marburg virus	viral protein genomic DNA.		
KW	Filovirus; viral protein; NP protein; VP30 protein; VP35 protein; VP40 protein; L protein; membrane-associated protein; matrix protein; polymerase complex protein; minor nucleoprotein; glycoprotein; GP; viral mutagenesis study; vaccine; gene therapy; gene; ds.			
OS	Marburg virus.			
PH	Key	Location/Qualifiers		
FT	CDS	103..2190		
FT		/*tag= a		
FT		/product= "NP protein"		
FT	CDS	2944..3933		
FT		/*tag= b		
FT		/product= "VP35 protein"		
FT	CDS	4567..5478		
FT		/*tag= c		
FT		/product= "VP40 protein"		
FT	CDS	5940..7985		
FT		/*tag= d		
FT		/product= "GP protein"		
FT	CDS	8867..9712		
FT		/*tag= e		
FT		/product= "VP30 protein"		
FT	CDS	10205..10966		

FT			/*tag= f	
FT			/product= "VP24 protein"	
FT	CDS		11479..18474	
FT			/*tag= g	
FT			/product= "L protein"	
XX				
FN			US2003215794-A1.	
XX				
PD			20-NOV-2003.	
XX				
PF			29-JAN-2003; 2003US-00353856.	
XX				
PR			31-JAN-2002; 2002US-0353972P.	
XX				
PA			(KAWA/) KAWAOKA Y.	
PA			(JASE/) JASENOSKY L D.	
XX			(NEUM/) NEUMANN G.	
PI			Kawaoka Y, Jasenosky LD, Neumann G;	
XX				
DR			WPI; 2004-010776/01.	
DR			P-PSDB; ADM48338, ADM48339, ADM48340, ADM48341, ADM48342, ADM48343, ADM48344.	
XX				
PT			Preparing filovirus useful in producing vaccines and gene therapy	
PT			vectors, comprises contacting a cell with a vector comprising a promoter	
PT			operably linked to a filovirus genomic cDNA or protein.	
XX				
PS			Disclosure; SEQ ID NO 27; 104pp; English.	
XX				
CC			The invention relates to a method of preparing filovirus. The method	
CC			involves contacting a cell with a vector comprising a promoter operably	
CC			linked to a filovirus genomic cDNA or its portion linked to a	
CC			transcription termination sequence or to a DNA segment encoding a	
CC			filovirus RNA transcriptase-polymerase, filovirus NP, VP30 or VP35, to	
CC			yield infectious filovirus. The portion of the cDNA when transcribed	
CC			yields a RNA capable of being packaged into filovirus virions or capable	
CC			of being replicated in the presence of filovirus proteins. The method is	
CC			useful for producing filovirus useful in viral mutagenesis studies and in	
CC			the production of vaccines and gene therapy vectors. The present sequence	
CC			is marburg virus (NP; L; polymerase complex-VP35; matrix protein-VP40;	
CC			glycoprotein-GP; minor nucleoprotein-VP30; membrane-associated protein-	
CC			VP24) genomic DNA.	
XX				
SQ			Sequence 19112 BP; 6176 A; 3757 C; 3578 G; 5601 T; 0 U; 0 Other;	
Query Match				
Best Local Similarity 72.0%; Score 1473.2; DB 13; Length 19112;				
Matches 1688; Conservative 0; Mismatches 358; Indels 0; Gaps 0;				
Qy	1	ATGAAGACCATATATTTTCTGATTAGTCTCAATTTTAAATCCAAAGTATATAAACTCTCCCT	60	
D	b	5940	ATGAAGACCATATCTCTTTATCAGTCTTATCTTAATCCAAAGGATATAAACTCTCCCT	5999
Qy	61	GTTTGTAGAAATGCTAGTAAACAGCAACCTCAAGATGATTCAGTGTGCTCCGGAAC	120	
D	b	6000	ATTTTAGAGATAGCTAGTAAACAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAC	6059
Qy	121	CTCCAAAGACAGAGATGTTTCATCTGATGGATTTACACAGAGTGGGCAAAAGTGTCT	180	
D	b	6060	CTCCAAAGACAGAGATGTTTCATCTGATGGATTTACACAGAGTGGGCAAAAGTGTCT	6119
Qy	181	GATTCCTCTTTGGAGCATCTTAAACGATGGCTTTTCAGGACAGAGTGTCTCTCCCAAGAAC	240	
D	b	6120	GATTCCTCTTTGGAGCATCTTAAACGATGGCTTTTCAGGACAGAGTGTCTCTCCCAAGAAC	6179
Qy	241	GTTGAGTATACGGAAGGAGAGAGCCAAAACATGTTTCAATATATAGTTACAGACCT	300	
D	b	6180	GTTGAGTATACGGAAGGAGAGAGCCAAAACATGTTTCAATATATAGTTACAGACCT	6239
Qy	301	TCTGGAATCTCTTGCTGCTGATCCCTCCAGTATATATCCGAGTATACCTTAATGTAAA	360	
D	b	6240	TCTGGAATCTCTTGCTGCTGATCCCTCCAGTATATATCCGAGTATACCTTAATGTAAA	6299

QY 361 ACTGTTTCATCATATATCAAGGTCAAAACCCCTCATGCAAGGGGATGGCCCTCCATTTGTGG 420
DB |||||
QY 6300 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCGCAAGGGATCGCCCTCCATTTGTGG 6359
DB |||||
QY 421 GGGGCAATTTTCTTGATGATCGGTTGCTCTCAACAATGTATACCGAGCAAGGTCCTTC 480
DB |||||
QY 6360 GGAGCATTTTCTTGATGATCGGTTGCTCTCAACAATGTATACCGAGGAGAGTCTTC 6419
DB |||||
QY 481 ACTGAAGGAATAATAGCAGCTATGATTTGTTAATAAGACAGTTTCACAGAATGATTTTCT 540
DB |||||
QY 6420 ACTGAAGGAATAATAGCAGCTATGATTTGTTAATAAGACAGTTTCACAGAATGATTTTCT 6479
DB |||||
QY 541 AGGCAAGGACAGGTTATCGTCAATGAACTTGACCTCCACCAATTAATATTTGACAGAAC 600
DB |||||
QY 6480 AGGCAAGGACAGGTTATCGTCAATGAACTTGACCTCCACCAATTAATATTTGACAGAAC 6539
DB |||||
QY 601 AGCAATGAAGCGCAGAGAAATGATACGGGATGTTTGGCATCTCTCCAAAGATACAACTCC 660
DB |||||
QY 6540 AACAAATGGAACAAACGAATGACACTGGATGCTTCGGTGCTCTTCAAGAAATACAACTCC 6599
DB |||||
QY 661 ACAAAATCAAAATGAGCCCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAACCTCCG 720
DB |||||
QY 6600 ACGAAGAAATCAAACTGCTCCGTCCTCCAAATACCCCTCACCACTGCCACAGCCCGTCCA 6659
DB |||||
QY 721 AGCAATCACTCTCAAAATACCTCAAAATTAATCTGCTAAATCTGGAACTATGAACCCAAAT 780
DB |||||
QY 6660 GAGATCAAAACCCACAGCACCCCACTGATGCCACCACTCACTCAACACCAACAGACCCCAAC 6719
DB |||||
QY 781 AGGCAAGTGAAGACCTTATGATTTCCGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
DB |||||
QY 6720 AATGATGATGAGGACCTCATAACTCCGTTTCAGGGTCCGGAGAAACAGGAACCCCTATACA 6779
DB |||||
QY 841 ACTCTTAATGTAGTCACTGAACAGAAACAAATCGTCAACAAATATTTGTCACCTCTCTCACTA 900
DB |||||
QY 6780 ACTTCAGTGGGTCACTAAGCAGGGGTTTCATCAACAAATGCCACCACTCCCTCACCAC 6839
DB |||||
QY 901 CATCAAGGACCTCACAACATGAGAGAAACAGTACGAATCTTCCCGACAGTGTGTAACT 960
DB |||||
QY 6840 CAACCAAGCAGCCACAGCAGAAAGAGAAACAAACACAGACCACTTCCCAAGGTACTGTGACT 6899
DB |||||
QY 961 GAGCACAATGGAACCGACCCACACACACACACCAACGCTCTCTCAACAACTACTAATACA 1020
DB |||||
QY 6900 GAACCCCAACAAACACACACACACACACCGTCCATGCCCCCCCCACACACCACTGCA 6959
DB |||||
QY 1021 ACTCCCACTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCCAACCCGCAAC 1080
DB |||||
QY 6960 ATCTCTACTAACAACCTCCAGAAACAACTTCAGCACCTCTCTGTATCACTACAAAC 7019
DB |||||
QY 1081 ATCAACCAATATGATACAAACGTGAATAGCAGAAAGCGCAACAAACCAATGTCTCAGTTG 1140
DB |||||
QY 7020 ACCACCAATTAAGCAGACAGCAGCAGCCACTGAAATGAAACAAACCAAGTGCCCTCG 7079
DB |||||
QY 1141 AACCAACTCTAGATCCACAGAAATCCACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1200
DB |||||
QY 7080 AAAACAAACCTGCTCCAAACAGGAAATCTTACCAAGCAGAAAGACAGCCTTAACACAGCAA 7139
DB |||||
QY 1201 AACATCATGAGCAGATCAGATATACAGAAACACACACCAACCAATTTCTTCCGGAT 1260
DB |||||
QY 7140 GGCCCCACCAACAGGCAACCAATATGACAAATGGGCATTTAAACAGTCCCTCCCCACC 7199
DB |||||
QY 1261 TCTAGTCCGACAAACCCGCCCTCTATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
DB |||||
QY 7200 CCCAACCCGACACACACATCTTGTAATTTTCAGAAAGAAACGAAGTATCTCTGGAGG 7259
DB |||||
QY 1321 GAAGCGCAGATGTTCCCTTTCTGATGGGTTAATAAATGCTCCAAATGATTTGACCCCA 1380
DB |||||
QY 7260 GAAGCGCAGATGTTTCTTTCTGACGGGTTAATAAATGCTCCAAATGATTTGATCCA 7319
DB |||||
QY 1381 GTTCCAAATACAAACCAATCTTTGATCAATCTCTAGTCTGCTGGCTCGGCTGAGGAA 1440
DB |||||
QY 7320 GTTCCAAATACAAACCAATCTTTGATCAATCTCTAGTCTGCTGGCTCGGCTGAGGAA 7379
DB |||||

QY 1441 GATCAACATGCTCCCCCATATATAGTTTAACTTTATCTTTATTTCCCTAATATAAATGAG 1500
DB |||||
QY 7380 GATCAACATGCTCCCCCATATATAGTTTAACTTTATCTTTATTTCTAATATAAATGAA 7439
DB |||||
QY 1501 AACACTGCTACTCTCGAGAGAAATGAGAAATGATGTGATGACAGAGTTAAAGAAATTTGGAGC 1560
DB |||||
QY 7440 AACACTGCTACTCTCGAGAGAAATGAGAAATGATGTGATGACAGAGTTAAAGAAATTTGGAGC 7499
DB |||||
QY 1561 GTTCAGAGGATGACCTGCGCCGAGGCTCAGTTGGATACCGTTTTTGGCCCTTGGAATTT 1620
DB |||||
QY 7500 GTTCAGAGGATGACCTGCGCCGAGGCTCAGTTGGATACCGTTTTTGGCCCTTGGAATTT 7559
DB |||||
QY 1621 GAAGGACTTTTACACTGCTGTTTTTAAATTAATAAATCABAACAAATTTGGCTCGCAGGTTGAGG 1680
DB |||||
QY 7560 GAAGGACTTTTATCTGCTGGTTTTAATTAATAAACCBAACAAATTTGGCTCGCAGGTTGAGG 7619
DB |||||
QY 1681 CGTCTAGCCCAATCAAACTGCAAACTCCTTGGAATCTTTATTTGAGAGTCACAACTGAGGAA 1740
DB |||||
QY 7620 CGTCTAGCCCAATCAAACTGCAAACTCCTTGGAATCTTTATTAAGAGTCACAACTGAGGAA 7679
DB |||||
QY 1741 AGAATCTCTCTTAATCAATAGACATGCTATTGATCTTTCTACTCAACAAGATGGGGAGGA 1800
DB |||||
QY 7680 AGGACATTTTCTTAATTAATAGACATGCTATTGATCTTTCTACTCAACAAGATGGGGAGGA 7739
DB |||||
QY 1801 ACATGCAAGTCTGTCACCTGATTTGTCATCGGATAGAGACTTGTCCAAAATATTT 1860
DB |||||
QY 7740 ACATGCAAGTCTGTCACCTGATTTGTCATTTGGAATAGAGACTTGTCCAAAATATTT 7799
DB |||||
QY 1861 TCAGAGCAAAATGACCAAAATTAATAAGAGCAACAAAGAGGGGACTGGTTGGGCTCG 1920
DB |||||
QY 7800 TCGGCAAAATGACCAAAATTAATAAGATGACAAAGAGGGGACTGGTTGGGCTCTA 7859
DB |||||
QY 1921 GGTGTTAAATGGTGACATCCGACTGGGTGTTCTTAACTTAACCTTGGGCAATTTGCTACTA 1980
DB |||||
QY 7860 GGTGTTAAATGGTGACATCCGACTGGGTGTTCTTAACTTAACCTTGGGCAATTTGCTACTA 7919
DB |||||
QY 1981 TTATCCATAGTGTCTTGATGCTCTATCTGATTTTGTGCTATCTTTACTAATATATC 2040
DB |||||
QY 7920 TTATCCATAGTGTCTTGATGCTCTATCTGATTTTGTGCTATCTTTTACCATAATATC 7979
DB |||||
QY 2041 GGATAA 2046
DB |||||
QY 7980 GGGTAA 7985
DB |||||
RESULT 11
ACC71551
ID ACC71551 standard; DNA; 7005 BP.
XX
AC ACC71551;
XX
XX
DT 10-JUL-2003 (first entry)
XX
DE VRC6702 (pVR1012-x/s Marburg GP (dTM)) plasmid.
XX
XX Virucide; vaccine; immune response; Ebola virus; Marburg virus;
KW Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;
KW viral infection; filovirus; circular; cyclic; ds.
XX
OS Synthetic.
XX
XX WO2003028632-A2.
XX
XX 10-APR-2003.
XX
XX 24-SEP-2002; 2002WO-US030251.
XX
XX 01-OCT-2001; 2001US-0326476P.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Nabel GJ, Yang Z, Sullivan N, Sanchez A;
XX
XX


```
DR WPI; 2003-371961/35.
XX New bimodal priming and boosting compositions, useful as viral vaccines,
PT specifically for eliciting an immune response against a filovirus or a
PT disease caused by infection with filovirus.
XX
PS Claim 1; Page 188-190; 219pp; English.
XX
CC The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection. This vector expresses the Marburg virus glycoprotein
CC without its transmembrane and intracellular domains
XX
SQ Sequence 7005 BP; 1931 A; 1758 C; 1588 G; 1728 T; 0 U; 0 Other;

Query Match 69.3%; Score 1418.8; DB 10; Length 7005;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 1618; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
DB 2034 ATGAAGACCATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 2093

QY 61 GTTTTAGAAAATGCTAGTAACAGCCAACTCTCAAGATGTAGATTTCAGTGTGCTCCGGAACT 120
DB 2094 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTGGTATGCTCCGGAAT 2153

QY 121 CTCGAAAGACAGAGATGTTTCATCTGATGGATTTCACATGAGTGGGCAAAAGTTGCT 180
DB 2154 CTCGAAAGACAGAGATGTTTCATCTGATGGATTTCACATGAGTGGGCAAAAGTTGCT 2213

QY 181 GATTCCCTTTTGGAGGATCTAAACGATGGGCTTTTCAGGACAGGTTGTTCTCCCAAGAAC 240
DB 2214 GATTCCCTTTTGGAGGATCTAAACGATGGGCTTTTCAGGACAGGTTGTTCTCCCAAGAAC 2273

QY 241 GTTGTAGTATACGGAAAGGAGAAAGCCAAACATGTTTCAATATAGTGTAAACAGACCT 300
DB 2274 GTTGTAGTATACAGAGGGGGAGGAGGAGCCAAACATGCTACATATAGTGTAAAGGATCCC 2333

QY 301 TCTGAAAAATCTTCTGCTGTGATCTCTCCAGTAATATCCGGGATACCCCTTAATATGTA 360
DB 2334 TCTGAAAAATCTTCTGCTGTGATCTCTCTCCAGTAATATCCGGGATACCCCTTAATATG 2393

QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGACAGAGGGATGTCCTCCATTTGG 420
DB 2394 ACTATCCATCATATTTCAAGGTCAAAACCTCATGACAGAGGGATGTCCTCCATTTATGG 2453

QY 421 GGGGCAATTTTCTGTATGATCGCTTGCCTCTACAAATGTATCCGAGGCAAGGTCTTC 480
DB 2454 GGAGCATTTTCTGTATGATCGCATTCCTCCCAACATGTATCCGAGGCAAGGTCTTC 2513

QY 481 ACTGAAGAAATATAGCAGCTATGATGTTTAAATAGACAGTTTCACAGATGATTTTTCT 540
DB 2514 ACTGAAGGAAACATAGCAGCTATGATGTTTAAATAGACAGTTTCACAGATGATTTTTCT 2573

QY 541 AGGCAAGGACAGGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTTGGACAGC 600
DB 2574 CGGCAAGGACAGGTTATCGTCACATGAACCTTGACCTTCATATATATTTGGACAGT 2633

QY 601 AGCAATGAACCGACAGAAATGATACGGGATGTTTGGCATCTCTCCAAAGATATCAACTCC 660
DB 601 AGCAATGAACCGACAGAAATGATACGGGATGTTTGGCATCTCTCTCCAAAGATATCAACTCC 660
```


421	Db	GGAGCATTTTTTCTGTATGATCGCAATTTGCCCTCCACAACAATGTATACGGAGGCAAGTCTTC	480
481	Qy	ACTGAAGGAAATATAGCAGCTATGATTTGTTAAATAAGACAGTTTCACAGAAATGATTTTTTCT	540
481	Db		
481	Db	ACTGAAGGAAATATAGCAGCTATGATTTGTTCAATAAGACAGTTGCACAAAATGATTTTCTCG	540
541	Qy	AGGCAAGGACAAAGTTATCGTCAATGAACTTTGACCTCCACCAATAAATATTTGGACAAGC	600
541	Db		
541	Db	CGGCAAGGACAAAGGTTACCGTCATATGAACTCTGACTTCTACTAATAAATAATTTGGACAAGT	600
601	Qy	AGCAATGAAACGCGACAGAGAAATGATACGGGATGTTTTGGCATCTCTCCAGAATAACAATCC	660
601	Db		
601	Db	AGTAAACGAGACGCAAAACGAATGACACTGGATGTTTTCGCGCTCTTCAAGAAATCAATTTCT	660
661	Qy	ACAAACAATCAAAATGCGCTCTCATCTCTTTAAACCTCCATCCCTGCCCAACAGATACTCCG	720
661	Db		
661	Db	ACAAAGAACCAAAACATGTGCTCCGTCCAAATACTCTCCACACTGCCCAACAGCCCGTCCG	720
721	Qy	AGCATTCCTCTACAAATATCTCAAAATTAATCTGCTAAATCTGGAACCTATGAACCCAAAGT	780
721	Db		
721	Db	GAGATCAAACTTCACAAGCACCCCACTGATGCCACCAACTCAATACGACGGACGCCAAGC	780
781	Qy	AGCGACGATGAGGACCTTATGATTTTCGGCTCTCAGGATCTGGAGAACAGGGGCCCCACACA	840
781	Db		
781	Db	AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAAACAGAAACCCCAACA	840
841	Qy	ACTCTTAATGTAGTCACTGAAACAGAAACNAATCGTCAACAATATTGTGCCACTCTCTCACTA	900
841	Db		
841	Db	ACTTCTGATGGGTTCACCMAGCAAGGGTTTCATCAACNAATGCCACCACTCCCTCAACA	900
901	Qy	CATCCAAAGCACTCACAACTAGGACAAAACAGTAGTACGAATCTCTCCGACATGCTGTAACT	960
901	Db		
901	Db	CAACCAAGCACGCCACAGCAGGAGGAGAAACAAACAAACCAATTTCCCAAGATGCTGTGACT	960
961	Qy	GAGCACANTGAAACGGACCCCAACAACAACACGACAGCTCTCTCAACNATCTAATAACA	1020
961	Db		
961	Db	GAATAGACAAAATAACAAACTGCAACACCGTCCATGCCCCCTCATAAACACTACACACA	1020
1021	Qy	ACTCCCACCTATAACACTCTCAAGTACAACTCTAGTACTCTCTCCCTCCAAACCCGCAAC	1080
1021	Db		
1021	Db	ATCTCTACTAACACACTCTCAACACNACTTTCAGCACTCTCTCTGCACCAATTCACAAAC	1080
1081	Qy	ATCACCAATAATGATATACAAACGTGAACTAGCAGAGAAACGCAAAACAAATGCTCAGTTG	1140
1081	Db		
1081	Db	ACCACCAATGACAAACACACAGACACAATCACTGAAATGAGCAAAACCAAGTGCCCCCTCG	1140
1141	Qy	AACACAACTTAGATCCACAGAAAAATCCGACCAACAGGACAGACACCAACAGCACAAACC	1200
1141	Db		
1141	Db	ATAACAACTTCCTCCAAACGGGAAATCCCAACACAGCAAAAGACACCCAGCAGCAAAAAA	1200
1201	Qy	AACATCATATGAGGACATCAGATATATCAAGAGAAACACCCCAACAATTTCTTCCCGAT	1260
1201	Db		
1201	Db	GGCCCCGCAACACGGCACCAACACGCAAAATGAGCAATTCACAGTCTCTCCCCCAACC	1260
1261	Qy	TCTAGTCCGACAAACCGGCCCTCTTATATCTATTAGAAAGAACGATCGATCCTCTGGAGG	1320
1261	Db		
1261	Db	CCCAGCTCGACTGCACACACTCTGTATATTTTCAGAGAAAGCGATCGATTTTCTGGAA	1320
1321	Qy	GAAGCGACATGTTTCCCTTTTCTGATGGGTAAATAAATGCTCCAAATGATTTTGACCCA	1380
1321	Db		
1321	Db	GAAGGTGATATATTTCCCTTTTTAGATGGGTAAATAAATACTGAAATTTGATTTTGATCCA	1380
1381	Qy	GTTCCAATACAAAACAACTTTTGATGAATCTCTAGTTCTGGTGCTCGCGCTGAGGAA	1440
1381	Db		
1381	Db	ATCCCAACACAGAAACAACTTTTGATGAATCTCCAGCTTTAATACTTCAACTAATGAG	1440
1441	Qy	GATCAACATGCGCTCCCCCAATAATTAGTTTAACTTTTATCTTATTTTCTTAATAATAAGT	1500
1441	Db		
1441	Db	GAAACAACACACTCCCCCCGAATATCAGTTTAACTTCTTATTTTCTGATATAAAATGGA	1500
1501	Qy	AACACTGCTACTCTGGAGAAAAATGAGAAATGATTTGTGATCAGAGTTTAAGAAATTTGGAC	1560
1501	Db		
1501	Db	GATACTGCTACTCTGGGGAAAAACAGAAATGATTTGATGACAGAGTTGAGATTTGGAGT	1560

Qy	1561	GTT	CAGAGAGATGACCTGGCGCAGGGCTCAGTTGGATACCGTTTTTTGGCCCTCGAATT	1620		
Db	1561	GTG	CAGGAGGACGATTTTGGCGCAGGGCTTAGCTGGATACCAATTTTGGCCCTCGAATC	1620		
Qy	1621	GA	GGAGCTTTTACACTGCTGTTTTAAATTTAAAAATCAAAACAATTTGGTCTGCAGTTTCGAGG	1680		
Db	1621	GA	GGAGCTCTATACTGCCGGTTTAATCAAAAATCAGAAACAATTTAGTTTGTAGGTTGAGG	1680		
Qy	1681	CGT	CTAGCCCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCAACAACCTGAGGAA	1740		
Db	1681	CGCT	TAGCTAATCAAACTGCTAAATCCTTGGAGCTCTTGTAAAGGGTCAACAACGAGGAA	1740		
Qy	1741	AGA	CAATTCCTCTTAATCAATAGACATGCTATTGAGACTTTCTACTCAACAAGATGGGAGGA	1800		
Db	1741	AGG	CAATTTTCTTAAATCAATAGGCATGCAATTTGACTTTTTTGTCTACGAGGTGGGGCGGA	1800		
Qy	1801	ACAT	GCAAACTGCTTGGACCTGATTTGTCATCGGGATAGAAGACTTTGTCCAAAAATATT	1860		
Db	1801	ACAT	GCAAGGTGCTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTTAAAAATATC	1860		
Qy	1861	TCAG	AGCAAAATTGACCAAAATTAATAAGGACGAAACAATAAGAGGAGACTGTTTGGGGTCTG	1920		
Db	1861	TCAG	AACAATTCGACAAAAATCAGAAAGGATGAACAATAAGGAGGAAACTGGCTGGGGTCTA	1920		
Qy	1921	GGT	GTTAAATGGTGGACATCGGACTGGGGTGTCTTACTTAACCTTGGGCATTTTCTACTA	1980		
Db	1921	GGT	GCAAAATGGTGGACATCTGACTGGGGTGTCTTCAACCAATTTGGGCATCTCTGCTACTA	1980		
Qy	1981	TTAT	CCCATAGCTGCTTTGATTGCTCTATCCTCTGATTTGTCTATCTTTTACTTAAATATATC	2040		
Db	1981	TTAT	CTATAGCTGTTCTGATTTGCTCTGCTCTGTATCTGTATCTGTCTGTAATCTTCACTAATAATACATT	2040		
Qy	2041	GGAT	ATAA 2046			
Db	2041	GGAT	ATAA 2046			
RESULT 14						
ID	ACC71562	standard; DNA; 6902 BP.				
XX	AC	ACC71562;				
XX	AC	ACC71562;				
DT	10-JUL-2003	(first entry)				
XX	DE	VRC6703(pVR1012x/s Marburgdelta TM/h (codon optimised)) plasmid.				
XX	KW	Virucide; vaccine; immune response; Ebola virus; Marburg virus;				
KW	KW	Lassa virus; retrovirus; paramyxovirus; influenza virus; adenovirus;				
KW	KW	viral infection; filovirus; circular; ds.				
OS	OS	Synthetic.				
PN	PN	WO2003028632 -A2.				
XX	PD	10-APR-2003.				
XX	PF	24-SEP-2002; 2002WO-US030251.				
XX	PR	01-OCT-2001; 2001US-0326476P.				
XX	PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.				
XX	PI	Nabel GJ, Yang Z, Sullivan N, Sanchez A;				
DR	DR	WPI; 2003-371961/35.				
XX	PT	New bimodal priming and boosting compositions, useful as viral vaccines,				
PT	PT	specifically for eliciting an immune response against a filovirus or a				
XX	XX	disease caused by infection with filovirus.				
PS	PS	Claim 1; Page 213-215; 219pp; English.				

New bimodal priming and boosting compositions, useful as viral vaccines, specifically for eliciting an immune response against a filovirus or a disease caused by infection with filovirus.

Claim 1; Page 213-215; 219pp; English.

XX The present invention relates to a bimodal priming composition and
CC boosting composition for priming and boosting an immune response to an
CC antigen in an individual. The compositions comprise (a) a priming
CC composition comprised of a DNA plasmid comprising a nucleic acid molecule
CC encoding Ebola, Marburg, Lassa, retrovirus, paramyxovirus or influenza
CC virus glycoprotein or nucleoprotein or their epitope-bearing domain, or a
CC DNA plasmid selected from ACC71521-ACC71563 and (b) a boosting
CC composition comprised of a replication-deficient adenovirus, comprising a
CC nucleic acid molecule encoding Ebola, Marburg, Lassa, retrovirus,
CC paramyxovirus, or influenza virus glycoprotein or nucleoprotein or
CC epitope-bearing domain, or a replication deficient adenovirus selected
CC from constructs ACC71521-ACC71563. The compositions are useful as viral
CC vaccines, specifically for eliciting an immune response against a
CC filovirus or a disease caused by infection with filovirus e.g. Ebola
CC virus infection
XX
SQ Sequence 6902 BP; 1753 A; 1978 C; 1707 G; 1464 T; 0 U; 0 Other;

Query Match 40.8%; Score 834; DB 10; Length 6902;
Best Local Similarity 64.3%; Pred. No. 1.3e-231;
Matches 1251; Conservative 0; Mismatches 695; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTCTGATAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT 60
DB |||||
QY 1923 ATGAAGACCATCTGCTTTCATCAGCCTGATCTCTGATCCAGGGCATCAAGACCCGTGCC 1982
DB |||||
QY 61 GTTTTAGAAATTGCTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTGCTCCGGAACC 120
DB |||||
QY 1983 ATCTGGAGATGCGCAGCAACACACGCCCCGAGAACGTTGGACGGGTGTCAGCGGCACC 2042
DB |||||
QY 121 CTCAAAAGACAGAAAGATGTTTCATCTGATGGGATTTACATGAGTGGGCAAAAAGTGTCT 180
DB |||||
QY 2043 CTGCAGAGACCGAGAGCTGCACCTGATGGCTTACCTGAGCGGCGCAGAGTGGCC 2102
DB |||||
QY 181 GATTCCTCTTTGGAAGCATCTAAACGATGGGCTTTTCAGGACAGAGTGTCTCCCAAGAAC 240
DB |||||
QY 2103 GACAGCCCTCTGGAGGCCAGCAAGAGTGGGCTTTTCAGGACCGGCGTCCGCCCAAGAAC 2162
DB |||||
QY 241 GTTGAGTATACGGAAGGAGAGAGCCAAACATGTTACATATTAAGTGTACAGACCCCT 300
DB |||||
QY 2163 GTGGAGTACACGAGGGCGAGGAGGCCAAGACCTCTGCTACATCATCAGCGTACCGACCC 2222
DB |||||
QY 301 TCTGAAAATCTTGCTGTGATCTCCCGAGTAAATATCCGGATATACCTCAAAATGTAAA 360
DB |||||
QY 2223 AGCGCAAGAGGCTGCTGTGACCTCCCAACATCAGGACTACCTTAAGTGCAAG 2282
DB |||||
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGCAAGGGGATGGCCCTCATTTGTGG 420
DB |||||
QY 2283 ACCATCCACCATCCAGGGCCAGAACCTCTACGCCAGGGCATCGCCCTGACCTGTGG 2342
DB |||||
QY 421 GGGCATTTTCTGTATGATCGCGTGGCTCTACCAACATGTACCGAGGCAAGGTCTTC 480
DB |||||
QY 2343 GCGCCTTCTTCTGTACAGGATCGCGAGCACCAATGTAACGGGCAAGGTGTTC 2402
DB |||||
QY 481 ACTGAAGAAATATACGAGCTATGATTTGTTAATAGACAGTTTCACAGAAATGATTTTTCT 540
DB |||||
QY 2403 ACCGAGGCAACATCGCGCCATGATCGTTAACAAGACCGTGCACAAGATGATCTTCAG 2462
DB |||||
QY 541 AGGCAAGGACAGGTTATGTCATATGAACTTTGACCTCCACCAATTAATTTGGACAAGC 600
DB |||||
QY 2463 AGGCAGGGCCAGGGCTACAGGCACATGAACCTTGACCAAGTACTTGGACCAAGC 2522
DB |||||
QY 601 AGCAATGAACGACAGAAATGATACGGGATGTTTGGCATCTCTCCAGAAATACAACTCC 660
DB |||||
QY 2523 AACACCGGACCCAGCAACCAACACCGGCTGCTTCGGGGCCCTCGAGAGGTACAAACAG 2582
DB |||||
QY 661 ACAAACAATCAAAATGCGCCTCCATCTCTTAAACCTTCAATCTCCCTGCCACAGTAACTCCG 720
DB |||||
QY 2583 ACCAAGACCGACCTGGGCCCCCAGCAAGATCCCGACGCCCTGCCCCACCGCAGGCC 2642
DB |||||
QY 721 AGCATTTCACTTACAAATACTCAAAATTAATATCTGCTAAATCTGGAACTATGAAACCAAGT 780
DB |||||

2643 GAGATCAAGCCCAACAGACCCCCCAACCGACCCCTGAAACACCAACGACGACCCCAAC 2702

781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840

2703 AACGACGACGAGGACCTGATCACAGCGGACGCGGCGGAGGAGGAGGAGGAGGAGGAG 2762

841 ACTCTTAATGATGACTGACGACAGAAACAATGCTCAACAATATTGTTCACTCTCTTCACTA 900

2763 ACCAGCGACGCGGTGACCAAGAGCGGCTGAGCAGCACCATGCTCTCTTACGCTAGCCCT 2822

901 CATCAAGGACCTCACAACAATGACAAACAGTACGAATCTTCCGACATGCTGTAAT 960

2823 CAGCCGACGACCTCTCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2882

961 GAGCACAAATGGAACCGACCCCAACACACAGCAGCAACGCTCTCAACAATACTAATAACA 1020

2883 GAGCCCAAGACCAACACACCGCCGAGCAGCATGCTCTCTCAACACACACGCGCC 2942

1021 ACTCCGACCTTAACACCTCTCAAGTACAACTCTAGTACTCTTCCCTTCCAAACCGGCAAC 1080

2943 ATCAGCACCAACACACGACGACGAGCAACATTTTACGACCTTGAGCGTGAGCGCTGAGAAC 3002

1081 ATCACCACATATGATACACACGTTGAATAGCAGAAAGCGAACAACCAATGCTCAGTTG 1140

3003 ACCACCACTACGACACCCGAGCAGCCGACCGGAGACGAGACGAGCAGCAGCGCCCTAGC 3062

1141 AACCAACTCTAGATCCAAACAGAAAATCCCAACAGGACAAAGACACCAACAGCAACAACC 1200

3063 AAGACCACTGCTCTCCACCGGCAACCTGACCAACCGCCAGAGACCAACCAACACCAAG 3122

1201 AACATCATGACGACATCATGATATTAACAGCAAAACACCCCAACAAATTTCTTCGGAT 1260

3123 GGCCCAACCAACCGCCCTTAAATGACCAACCGCCCACTGACGACGCGCCCAACGCGCCACC 3182

1261 TCTAGTCGACAAACCGCCCTCTATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320

3183 CCCAACCCCAACCGACCTGCTGTTACTTTCAGGAAAGAGAGGAGCATCTCTGGAGG 3242

1321 GAAGCGACATGTTTCTGATGGGTTTAAATGCTCCAATTTGATTTTGACCCA 1380

3243 GAGGCGATGTTTCTCTGACGCGCTGATCAACGCGCTTATCGACTTCGACCCC 3302

1381 GTTCCAAATACAAAACAATCTTTGATGAATCTCTCTAGTCTGTTGCTCTCGCTGAGGAA 1440

3303 GTGCCCAACACCAAGACCATCTTCGACGAGACGACGACGCGCGCGCCGCGGAGGAG 3362

1441 GATCAACATGCTCCCAATATTAGTTTAACTTTATCTTATTTTCTTAATAATAATAATGAG 1500

3363 GACGACGCGCAGCCCAACATCAGCCTGACCTGAGCTACTTCCCAACATCAACGAG 3422

1501 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGATGACAGATTAAGAAATTTGGAGC 1560

3423 AACCGCCTACAGCGCGGAGAACGAGAACGACTGCGACGCGGAGCTGAGGATCTGGAGC 3482

1561 GTTCAGGAGGATGACCTGGCGCGAGGCTCAGTTGGATACCGTTTTTTTGGCCCTCGAATT 1620

3483 GTGACGAGGAGGACGACCTGGCGCGCGCTGAGCTGGATTCTCTTCTCGGCGCCGCGCATC 3542

1621 GAAAGATTATACATGCTGTTTTTAATAAATAACAAACAAATTTGGTCTGCAAGTTGAGG 1680

3543 GAGGCGCTGTACACCGCGCGCTGATCAAGAAACAGAAACAACTGGTGTGACGGCTGAGG 3602

1681 CGTCTAGCAATCAAACTGCAAAATCTTTGGAATCTTTATTTGAGAGTCAACAACTGAGGAA 1740

3603 AGGTGCGCAACAGACCGCCCAAGCTGAGCTGCTGAGGGTGCTGAGGAGGAGGAGGAGGAG 3662

1741 AGAACAATCTCTTAATCAATAGACATGCTATTGATTTTCTCACTCAAGATGGGAGGA 1800

3663 AGGACCTTCAGCCTGATCAACAGGACGCGCATTCGACTTCTGCTGACAGGTTGGGCGCGC 3722

1801 ACATGCAAGTCTTGGACCTGATTTGTTGATCGGATAGAGACTTGTCCAAAAATATT 1860

3723 ACCTGCAAGGTGCTGGGCGCCCGACTGCTGCTATCGGATCGAGGACCTGAGCAGGAAACATC 3782

Db 1261 CCCAGCTCGACTGCGACAAACATCTTGTATATTTCAGAGAAAGCGATCG 1308

Search completed: July 31, 2005, 16:57:29
Job time : 1077.87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:13:05 ; Search time 6679.18 Seconds
(without alignments)
11660.032 Million cell updates/sec

Title: US-10-066-506A-7
Perfect score: 2046
Sequence: 1 atgaagaccatattttct.....ttactaaatatcggataa 2046

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	85.6	4.2	895	9	CNS0071A
C 2	80.8	3.9	1101	9	CNS00FXE
C 3	80.4	3.9	922	9	CNS0073W
C 4	79.8	3.9	909	9	CNS007TL
C 5	74.8	3.7	884	9	CNS006UO
C 6	72	3.5	1101	9	CNS00LOO
C 7	69	3.4	1101	9	CNS010OX
C 8	67.4	3.3	1225	9	CNS016KK
C 9	66.2	3.2	1101	9	CNS017ZT
C 10	66	3.2	1101	9	CNS00LT2
C 11	65.8	3.2	1101	9	CNS016K3
C 12	65.4	3.2	769	8	B2579780
C 13	64.4	3.1	1797	9	AG390561
C 14	64.2	3.1	1179	9	AG117870
C 15	64	3.1	1052	6	CD389522
C 16	63.8	3.1	939	9	CNS00CNG
C 17	63.6	3.1	860	9	CNS018FL
C 18	63	3.0	937	9	CNS006ST
C 19	62.4	3.0	861	9	CNS0075A
C 20	62.2	3.0	935	9	AG127540
C 21	61.6	3.0	1037	9	AG379951
C 22	61	3.0	1325	9	CG754898
C 23	60.8	3.0	905	9	AG129358
C 24	60.6	3.0	926	8	AQ782232

25	60.4	3.0	1095	9	CNS04428
26	59.6	2.9	902	9	CNS006QP
27	59.6	2.9	1052	9	CNS00DIP
28	59.4	2.9	1101	9	CNS017WI
29	59.4	2.9	1188	9	CG747780
30	59.2	2.9	1974	9	CL467385
31	58.6	2.9	1257	9	AG289995
32	58.6	2.9	1499	9	AG371052
33	58.4	2.9	1233	8	BZ696963
34	58.4	2.9	1779	9	AG390408
35	58	2.8	891	5	BU530351
36	57.8	2.8	1713	9	CG754301
37	57.6	2.8	779	9	AG519641
38	57.4	2.8	768	9	AG599644
39	57.4	2.8	1461	9	CL647640
40	57.4	2.8	2050	9	AG365337
41	57.2	2.8	1226	9	CG757449
42	57	2.8	1144	9	AG126238
43	57	2.8	1185	5	BQ650581
44	57	2.8	1759	9	CG754278
45	56.8	2.8	513	9	CNS02CMP

ALIGNMENTS

RESULT 1
CNS0071A/c
LOCUS
DEFINITION
CNS0071A 895 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit fly), Genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AL066286
GI:4945153

GSS.
Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 895)

Genoscope.

Direct Submission

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mammose in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

ECORI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain Y2; cn bw sp, the same strain used for the BDGP's

P1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers
1..895
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14B09"
/note="end : TET3"

ORIGIN

Query Match 4.2%; Score 85.6; DB 9; Length 895;
Best Local Similarity 22.5%; Pred. No. 1.9e-12;

```
Matches 96; Conservative 165; Mismatches 164; Indels 1; Gaps 1;
QY 855 CACTGAACAGAAACAATCGTCAACAATATTGTCCACTCTTCACTATCATCATCAAGCACCTC 914
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
889 CCCMMACCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 830
QY 915 ACAACATGAGCAAAACAGTACGAATCTCTCCGACATCTGTAACTGAGCACAATGGAAAC 974
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
829 MMAMMCCCMCMCMCAAAAMCMAMMMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 770
QY 975 CGACCAACAACAACAGCAACGCTCTCTCAACATACTAATACAACTCCACCTATAA 1034
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
769 MMAMCMAMMAAMAMCMAMMAAMAMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 710
QY 1035 CACTCTCAAGTACAACCTCAGTACTCTCTCCCTCCAAACCGCAACATCACCAATAATGA 1094
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
709 MCAAMAMAMAMAMAMAMAMAMAMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 650
QY 1095 TACACAACGTGAATGAGAAAGCAACAAACCAATCTCAGTTGAAACAACATCTCTAGA 1154
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
649 MAMCCAMMCAAMM-AMCMCMCMCCCAAMCMCAACCCMMCMCMAMAMAMAMCMCMCMCM 591
QY 1155 TCCACAGAAAATCCCAACAGGACAGACCAACCAACCAACATCATCATGAC 1214
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
590 MMAMCMCMAMMCMAMMAMMAACAAMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 531
QY 1215 GACATCAGATATAACAAGCAACACCCACCAAAATTTCTTCGGATTCTAGTCCGACAA 1274
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
530 MMAMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 471
QY 1275 CGGCC 1280
Db |:::|
470 CACCM 465

RESULT 2
CNS00FXE/c 1101 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC:
DEFINITION BACR32C19 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL071370.1 GI:4951210
VERSION AL071370.1
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Drosophila melanogaster (fruit fly)
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
location/Qualifiers
1. .1101
/organism="Drosophila melanogaster"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR32C19"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 3.9%; Score 80.8; DB 9; Length 1101;
Best Local Similarity 25.7%; Pred. No. 4.5e-11;
Matches 105; Conservative 130; Mismatches 174; Indels 0; Gaps 0;

QY 906 AAGCACCCTCAACATGAGCAAAACAGTACGAATCTCTCCGACATCTGTAACTGACGA 965
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1101 AAAMAMATAAMMAHAYACAAAAMAMCMMMMAACAAMAMCMMAAMCMCMCMMAA 1042
QY 966 CAATGGAAACGAGCCCAACACACACAGCAACGCTCTCAACAATACTATAACAATCC 1025
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
1041 MAMMAMCAGMHCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 982
QY 1026 CACCTATAACACTCTCAAGTACAAACCTCAGTACTCTCTCCCTCCAAACCGCAACATCAC 1085
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
981 MAMCMAMMMMNCAAMCMMAAMAMAMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 922
QY 1086 CAATAATGATACACAACGTGAATAGAGAAAGGAAACAACCAATGCTCAGTTGAAACAC 1145
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
921 MACCCATACAMMTTMMCMMAACMAMAMMAAAAMAMAMAMAMAMAMAMAMAMAMAM 862
QY 1146 AACTCTAGATCCCAAGAAATCCCAACAGGACAGACCAACCAACAGCAACCAACAT 1205
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
861 AAMAMAMAMAMAMAMAMAMAMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 802
QY 1206 CATCATGACGACATCAGATATAACAAGCAACACCCCAACAAATTTCTTCGGGATTTCTAG 1265
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
801 MCCAMMAMCMCMCSAGMAMCMMAACMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 742
QY 1266 TCGACAAACCGCCCTCTATATATCTTTAGAAAGAAAGATCGATCTCTC 1314
Db |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
741 CMCMAMACCCCMCAACCCCTCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 693

RESULT 3
CNS0073W/c 922 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL066784
VERSION AL066784.1 GI:4945247
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 922)
AUTHORS Direct Submission
TITLE Drosophila melanogaster (fruit fly)
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
location/Qualifiers
1. .1101
/organism="Drosophila melanogaster"
```


[illegible]

Qy	1161	AGAAATCCACACACAGGACACACACACACACACCAACCAA	1202
Dd	413	CACACNCAYMCACCCCWYAATAAAMAHMAAHMAHACRAAA	372
RESULT 8			
CNS0166K/c			
LOCUS			
DEFINITION	CNS0166K Drosophila melanogaster genome survey sequence SP6 end of BAC BACN15C20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.		
ACCESSION	AL106358		
VERSION	AL106358.1 GI:5621512		
KEYWORDS	GSS.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 1225) Genoscope.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - web : www.genoscope.cns.fr)		
JOURNAL	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.		
FEATURES	Location/Qualifiers		
source	1..1225 /organism="Drosophila melanogaster" /mol_type="genomic DNA" /db_xref="taxon:7227" /clone="BACN15C20" /clone_lib="DrosBAC" /plasmid="pBelobAC11" /note="end : SP6"		
ORIGIN			
Query Match	3.3%; Score 67.4; DB 9; Length 1225;		
Best Local Similarity	32.3%; Pred. No. 2.9e-07;		
Matches 136; Conservative	97; Mismatches 191; Indels 1; Gaps 1;		
Qy	822	AGAACAGGGGCCCACACAACTCTTAATGTAGTGCATCTGAACAGAAAACAATCGTCAACAAT	881
Dd	1127	ACWACAWACCCCMWACMACCCACACACACMMWCACAWACMCMCWAACCC 1068	
Qy	882	ATTGTCCACTCTTTCTACTCATCCAAGCACTTCACAAATGAGCAAAACA-GTACGAATC	940
Dd	1067	MWAAWMCMCHCMCMCCACCMCMCMCCACAAWAMCAAATAAAAMCACAACMCMCAA	1008
Qy	941	CTTCCCGACATGTGTGTAACGTAGCAGCAATGGACCACCAACACACACACACCAACGC	1000
Dd	1007	MACACCCCMWCCCAAAAAACCCACCCCAAACACMCMCAACACMCMCCCCCCCCCMCM	948
Qy	1001	TCTCTCAACAATAGTATACAACTCCCACTTATACACTCTCAAGTACAACTCTAGTACT	1060
Dd	947	ACMCMCAAMMMCAACCCCMWMAACCCCAACMMMMMMCCCMCMCMCCAMMMCMCMCM	888
Qy	1061	CTTCCCTCCCAACCCGCAACATCACCACTAATGATACAAACGTGAACTAGCAGAAAGCG	1120
Dd	887	CAACMMWCMCHCMCCMWCMWCMWCMWCAAAWCAACAAACACACACMMAACWACCCC	828
Qy	1121	AACAAACCAATGTCAGTTGAAACAACAATCTAGATCCAAACAGAAAAATCCACACACAGGAC	1180
Dd	827	CCCCCCCCCCCCCCCCCMWCCCAACCMMAAAWAAACAAAAAAAACCCMCCACAAW	768


```
QY 1181 AAGACACCAAGCAGCACCACCATCATCATGACGATCATATATAACAGCAACACC 1240
Db 767 ACCCCCCHCCCCCMAAAMAAAATKTMAMAAACAACAAMAAAYCMCCGRRARVMSSS 708
QY 1241 CCACAAA 1247
Db 707 GVSVVV 701

RESULT 9
CNS017ZT/c
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC
BACN37N10 of Drosophila library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL108707
VERSION AL108707.1 GI:5629011
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (BDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CDPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.

FEATURES
source Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN37N10"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : T7"

ORIGIN
Query Match 3.2%; Score 66.2; DB 9; Length 1101;
Best Local Similarity 31.4%; Pred. No. 6.3e-07;
Matches 143; Conservative 113; Mismatches 194; Indels 6; Gaps 2;

QY 887 CCATCCTTCACTACATCCAGCACCTCACACATGACGAAACAGTAGCATCTTCCC 946
Db 1095 CCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCC 1036
QY 947 GACATGCTGTAACTGAGCAGCATGGAACCGACCAACAAACACACAGCATGCTCTCA 1006
Db 1035 CAAACACACACACACACACACACACACACACACACACACACACACACACACAC 976
QY 1007 ACAATACTAATAACAATCCCACTTCAAGTCAACACTCTCAAGTCAACACTCTCTCCC 1066
Db 975 MWACACCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCCMCC 916
QY 1067 CTCGACCCGACATGACCAATATGATACACACGTGAACATGACGAGAAACGACAA 1126
Db 915 AMCCAAMACACACACACACACACACACACACACACACACACACACACACACAC 860
QY 1127 CCAATGCTCAGTTGAACACAACTCTAGATCCAAAGAAAATCCCAACACAGACAGA 1186
Db 859 CCACACACACACACACACACACACACACACACACACACACACACACACACAC 800
QY 1187 CCAACAGCACACACACACACATCATGACGATCATGATATATAACAGCA--AACACCC 1244
```

```
Db 799 CCMCACAACMCMYAMCMCAWCHYCTTTTTTTTTTNTACWYCMAAAHMAYAYATTACTYC 740
QY 1245 AAAATCTTCTCCGATTTCTAGTCGCAACACCGCCCTCTATATACTTTTGAAGAAACG 1304
Db 739 AAAMWYAYYYNYMYCTCTHMYATACYMCCYAYHYWYTWYTWCMCTWHAMWMTWY 680
QY 1305 ATCATCTCTGAGGAGGAGGACATGTTCCCTTT 1340
Db 679 TYATMCMCAAMCTMMCAAWTCATCMCCCTCCCTTY 644

RESULT 10
CNS00LT2/c
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC;
BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL078714
VERSION AL078714.1 GI:5102004
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Direct Submission
TITLE Submitted (11-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department
of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR48P19"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 3.2%; Score 66; DB 9; Length 1101;
Best Local Similarity 13.1%; Pred. No. 7.1e-07;
Matches 53; Conservative 194; Mismatches 157; Indels 0; Gaps 0;

QY 835 CACACAACTCTTAATGTAGTCATCGAAGAGAAACAATCGTCAACAATATTGTCCACTCT 894
Db 1092 MMYMMCMCHHTAAMMMAMMMMMMMMMMMMMMMMMMMCMCMCMCMCMCMCMCMCHMM 1033
QY 895 TCATCATCTCAACGACCTCACACATGACGAGAAACAGTAGCATCTCTCCGACATGCT 954
Db 1032 MMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM 973
QY 955 GTAACGTGACGACAACTGGAACCGACCAACACACACAGCAACGCTCTCTCAACAATACT 1014
Db 972 TTYMMAMCMTHTHYMMMMMMMMMMMMMMMMMMMMMMCMCMCMCMCMCMCMCMCMCM 913
```

```

Qy 1015 AATCAACTCCCACTATAACACTCTCAAGTACCAACTCAGTACTCTCTTCCCTCCCAAC 1074
Db 912 TTHHHMMCCMCMCCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 853
Qy 1075 GCGCAATCACCATAATATGATACACAACGCTGAAGTACGAGAAAGCGAACCAACATGCT 1134
Db 852 MWMWCMCHMAMMTTTTTTTTTTHMMAMMAVHTTMMMTTMMWMMCMWMMWMAAATTTMMWMM 793
Qy 1135 CAGTTGACACAACATCTAGATCCACAGAAAATCCACCACGACAGACACCAACAGC 1194
Db 792 AMHTHHCTMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 733
Qy 1195 ACAACCAACATCATGACGACATCAGATATACAAGCAACA 1238
Db 732 HMMAMATMTCHMMTHMMHTTTTMMMMMAAAWTTTWAHA 689

RESULT 11
CNS016K3/1101 bp DNA linear GSS 26-JUL-1999
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC
BACN16A14 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL106845
VERSION
AL106845.1 GI:5624103
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (23-JUL-1999) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
JOURNAL
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN16A14"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : T7"

ORIGIN
Query Match 3.2%; Score 65.8; DB 9; Length 1101;
Best Local Similarity 28.5%; Pred. No. 8.1e-07;
Matches 131; Conservative 138; Mismatches 185; Indels 6; Gaps 2;

Qy 832 CCCACACAACCTCTAATGCTAGTACTGACGAGAAACATCGTCAACATATGCTCCACT 891
Db 1086 YHHHHYAAAMHHMMHMYAAHCHYAAAYYYHHAAHYYYHYHHYHYHYHYHYHYHY 1027
Qy 892 CTTCACTACATCAAGCACTCAACATGACGAAACAGTACGAATCTCTCCGACAT 951
Db 1026 YHMHWHYHHHHAAHAYYWAAYYYATTAATAAATAAATAAACHAACHNAAYY 967
Qy 952 GCTGTAACTGACCAATGGAACCGACCAACAACACAGCAACGCTCTCTCAACAT 1011
Db 966 -----AACVHYACMHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHA 912
Qy 1012 ACTAATACAACCTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCA 1071

```

```

Db 911 HYYTAAAYAAAYYYYHYAAAMCHYHMYAAHAAACACYYYHYAAAHYACYACTAAAYY 852
Qy 1072 ACCGCAACATCACCATAATATGATACACAACGCTGAAGTACGAGAAAGCGAACCAACAT 1131
Db 851 MHAHAHHYHYHYAAAYHYAAAHCHYCHYCHYCHYCHYCHYCHYCHYCHYCHYCHYCHY 792
Qy 1132 GCTCAGTTGACACAACATCTAGATCCACAGAAAATCCACCACGACAGACACCAAC 1191
Db 791 AYACAAAACAAACAMTAAHAAHAAACHYCCACAACHCCCMC-CCACAACAACCCACM 733
Qy 1192 AGCAACCAACCAATCATCATGACGACATCAGATATAACAAGCAACACCCCAACATCT 1251
Db 732 ACCMMCCCMCMYHYHAAACMCAACACBACABMACAYMTMYCYCAACMMAYCC 673
Qy 1252 TCTCCGATTTCTAGTCCGCAACACCGCCCTCTATATATACT 1291
Db 672 ACCMCAVACCCCAACMBBACWACBCTCTCNNNNNNNCK 633

RESULT 12
BZ579780 769 bp DNA linear GSS 17-DEC-2002
LOCUS
DEFINITION
msh2_769.y2 msh Pseudomonas aeruginosa genomic clone msh2_769,
genomic survey sequence.
ACCESSION
BZ579780
VERSION
BZ579780.1 GI:27214841
KEYWORDS
GSS.
SOURCE
Pseudomonas aeruginosa
ORGANISM
Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE
1 (bases 1 to 769)
Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
COMMENT
Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 20622216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.
FEATURES
source
Location/Qualifiers
1..769
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="MSH"
/db_xref="taxon:287"
/clone="msh2_769"
/clone_lib="msh"
/note="Environmental isolate. Whole genomic shotgun
library."

ORIGIN
Query Match 3.2%; Score 65.4; DB 8; Length 769;
Best Local Similarity 49.6%; Pred. No. 9.6e-07;
Matches 195; Conservative 0; Mismatches 196; Indels 2; Gaps 1;

Qy 855 CACTGAACAGAAAACATCGTCAACAATATTGTCCACTCTCTTCACTACATCAAGCACTC 914
Db 173 CACCAAAAACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 232
Qy 915 ACAACATGAGAAAACAGTAGGAATCTTCCGACATCTGCTTAAGGACATGGAC 974
Db 233 ACAACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 292
Qy 975 CGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1034
Db 293 CACCACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 350

```

```
QY 1035 CACTTCTCAAGTACACCTCAGTACTCTTCCCTCCACCGCAACATCACCATAATGA 1094
|||
Db 351 CACAACAAACACACACAAACACACCTCAACACCAACACACCAACACCAACCCACCC 410
|||
QY 1095 TACACAAGCTGAATAGCAGAGAAAGGAACAAACCAATGCTCAGTTGAACACACTCTAGA 1154
|||
Db 411 AACACACACACACGAGAAACAAAAAACACCAACAAAAACCAACCAACCCACCCCA 470
|||
QY 1155 TCCACAGAAAATCCACACACAGACAGACACCAACAGCAGCAACCAACATCATCATGAC 1214
|||
Db 471 CACACACAAACCAACACACACCCACCAACACACCAACCAACCAACCAACCAACAA 530
|||
QY 1215 GACATCAGATATAACAAGCAACACCCCAACAA 1247
|||
Db 531 AAAACATAACACACACACACGACGACCA 563
|||

RESULT 13
AG390561/c
LOCUS Mus musculus molossinus DNA, clone:MSMg01-207B07.TJ, genomic survey
DEFINITION
ACCESSION AG390561
VERSION AG390561.1 GI:48007799
KEYWORDS GSS.
SOURCE Mus musculus molossinus
ORGANISM Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1797)
Direct Submission
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBac3.6
R.Site 1 : EcoRI.
R.Site 2 : EcoRI.
FEATURES
source
1..1797
Location/Qualifiers
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-207B07.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN
Query Match 3.1%; Score 64.4; DB 9; Length 1797;
Best Local Similarity 47.8%; Pred. No. 2.3e-06;
Matches 196; Conservative 0; Mismatches 213; Indels 1; Gaps 1;

QY 822 AGAACAGGGCCCAACACACTCTTAATGTAGTACTCTGAACAGAAACAAATCGTCAACAA 881
|||
```

```
Db 1531 AANACAAACACACAAACAAATCAAAACACACACACACACAAACNCACACAAACCA 1472
|||
QY 882 ATTGTCACCTCTTCACTACATCAAGCAGCAGCTCTCAACATGAGCAAAACAGTAGCAATCC 941
|||
Db 1471 ACNACCAACACACACACACACACACACACACACACACACACACACACACACACAA 1412
|||
QY 942 TTCCCGACATGCTGTAATGAGCAATGGAGCCGACCCCAACACACACACAGGACGCT 1001
|||
Db 1411 CACNACCAACACACACACACACACACACACACACACACACACACACACACACACAA 1352
|||
QY 1002 CCTCAACAAATCAATCAACTCCCACTTAACTCACTCTCAAGTACACCTCAGTACTCC 1061
|||
Db 1351 CCACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1292
|||
QY 1062 TTCCCTCCAAACCGCAACATCACTAATATGATACACAAACAGTGAATCTAGCAGAAAGCA 1121
|||
Db 1291 AACACACACACACACAA-AACACACAAACAAACAAACAAACACACACACACACAA 1233
|||
QY 1122 ACAACCAATGCTCAGTTGAACACAACTCTAGATCCACAGAAATCCCAACACAGGACA 1181
|||
Db 1232 AAACACACACACACACACACACACACACACACACACACACACACACACACACAA 1173
|||
QY 1182 AGACACCAACAGCAGCAACCAACATCATCATGACGACATCAGATATAACAA 1231
|||
Db 1172 CAACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1123
|||

RESULT 14
AG117870/c
LOCUS Pan troglodytes DNA, clone: PTB-125N11.R, genomic survey sequence.
DEFINITION
ACCESSION AG117870
VERSION AG117870.1 GI:16738389
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 1179)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimpbes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY
Vector : pKS145
R.Site 1 : SacI.
R.Site 2 : SacI.
FEATURES
source
1..1179
Location/Qualifiers
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-125N11.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match 3.1%; Score 64.2; DB 9; Length 1179;
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 15:35:29 ; Search time 327.71 Seconds
(without alignments)
10215.808 Million cell updates/sec

Title: US-10-066-506A-7
Perfect score: 2046
Sequence: 1 atgaagaccatattttct.....ttactaaatatatcggtataa 2046

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1822	89.1	2247	3	US-08-760-615-5
2	1514.8	74.0	2164	3	US-08-760-615-3
3	1514.8	74.0	11460	4	US-09-336-910A-1
4	134	6.5	2172	3	US-08-760-615-1
5	134	6.5	2298	4	US-09-650-086A-1
6	54	2.6	5163	3	US-08-700-651-1
7	54	2.6	5163	3	US-08-928-361B-4
8	54	2.6	5163	4	US-09-588-995A-4
9	54	2.6	5318	3	US-08-700-651-2
10	54	2.6	5318	4	US-08-928-361B-3
11	54	2.6	5318	4	US-09-588-995A-3
12	48.8	2.4	631	4	US-09-270-767-2848
13	48.8	2.4	631	4	US-09-270-767-18130
14	48	2.3	5511	3	US-08-928-361B-2
15	48	2.3	5511	4	US-09-588-995A-2
16	48	2.3	7334	3	US-08-928-361B-1
17	48	2.3	7334	4	US-09-588-995A-1
18	47	2.3	767677	4	US-09-949-016-12147
19	47	2.3	767677	4	US-09-949-016-17361
20	43.6	2.1	832	4	US-09-621-976-2813
21	43.4	2.1	1141	4	US-09-806-708B-22
22	43.2	2.1	700	3	US-09-236-097-11
23	43	2.1	152132	4	US-09-949-016-13845
24	43	2.1	152145	4	US-09-949-016-12371
25	43	2.1	325034	4	US-09-949-016-14957
26	43	2.1	389504	4	US-09-949-016-11774
27	42.4	2.1	49559	4	US-09-949-016-12892

28	42.4	2.1	49559	4	US-09-949-016-17267	Sequence 17267, A
29	41.4	2.0	1968	4	US-09-248-796A-1729	Sequence 1729, Ap
30	40.8	2.0	7218	1	US-08-432-463-14	Sequence 14, Appl
31	40.8	2.0	96987	4	US-09-949-016-14429	Sequence 14429, A
32	40.6	2.0	55195	4	US-09-949-016-15854	Sequence 15854, A
33	40.4	2.0	582	4	US-09-787-292-3	Sequence 3, Appli
34	40.2	2.0	601	4	US-09-949-016-21189	Sequence 21189, A
35	40.2	2.0	601	4	US-09-949-016-158630	Sequence 158630, A
36	40.2	2.0	1086	1	US-08-415-751-27	Sequence 27, Appl
37	40.2	2.0	1086	1	US-08-415-751-28	Sequence 28, Appl
38	40.2	2.0	17723	4	US-09-949-016-15832	Sequence 15832, A
39	39.8	1.9	1397	4	US-09-216-393B-343	Sequence 343, App
40	39.8	1.9	1397	4	US-09-216-393B-345	Sequence 345, App
41	39.8	1.9	60990	4	US-09-949-016-14080	Sequence 14080, A
42	39.6	1.9	867	4	US-09-216-393B-340	Sequence 340, App
43	39.6	1.9	867	4	US-09-216-393B-342	Sequence 342, App
44	39.6	1.9	938	3	US-09-078-294-16	Sequence 16, Appl
45	39.4	1.9	601	4	US-09-949-016-144692	Sequence 144692, A

ALIGNMENTS

RESULT 1
US-08-760-615-5
; Sequence 5, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2247 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Ravn
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97..2142
US-08-760-615-5

Query Match 89.1%; Score 1822; DB 3; Length 2247; Best Local Similarity 93.2%; Pred. No. 0; Matches 1906; Conservative 0; Mismatches 140; Indels 0; Gaps 0;									
Qy	1	ATGAAGACCATATATTTCTGATTAGTCTCTATTTTAAATCCAAAGTATATAAAACTCTCCCT	60						
Db	97	ATGAAGACCATATATTTCTGATTAGTCTCTATTTTAAATCCAAAGTATATAAAACTCTCCCT	156						
Qy	61	GTTTTAGAAATTTGCTAGTAACAGCCCAACTCAAGATGTAGATTCAAGTGTCTCCGGAACC	120						
Db	157	GTTTTAGAAATTTGCTAGTAACAGCCCAACTCAAGATGTAGATTCAAGTGTCTCCGGAACC	216						
Qy	121	CTCCAAAAGACAGAGATGTTTCATCTGATGGATTTTACATGAGTGGCCAAAAGTTGCT	180						
Db	217	CTCCAAAAGACAGAGATGTTTCATCTGATGGATTTTACATGAGTGGCCAAAAGTTGCT	276						
Qy	181	GATTTCCCTTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC	240						
Db	277	GATTTCCCTTTTGGAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC	336						
Qy	241	GTTGAGTATACGGAAGGAGAAAGCCAAACATGTTTACAAATATAAGTGTAAACAGACCT	300						
Db	337	GTTGAGTATACGGAAGGAGAAAGCCAAACATGTTTACAAATATAAGTGTAAACAGACCT	396						
Qy	301	TCTGGAATCTTGTCTGTGATCTCTCCAGTAAATATCCGCAATTACCCTAATGTAAA	360						
Db	397	TCTGGAATCTTGTCTGTGATCTCTCCAGTAAATATCCGCAATTACCCTAATGTAAA	456						
Qy	361	ACTGTTTCATCATATTCAGGTCMAAACCTCATGACAGGGGATGGCCCTCCATTTGG	420						
Db	457	ACTGTTTCATCATATTCAGGTCMAAACCTCATGACAGGGGATGGCCCTCCATTTGG	516						
Qy	421	GGGGCATTTTCTGTATGATCGCGTGCCTCTACAAATATGTACCGAGGCAAGTCTTC	480						
Db	517	GGGGCATTTTCTGTATGATCGCGTGCCTCTACAAATATGTACCGAGGCAAGTCTTC	576						
Qy	481	ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAGACAGTTCACAGAAATGATTTTCT	540						
Db	577	ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAGACAGTTCACAGAAATGATTTTCT	636						
Qy	541	AGGCAAGGACAGGTTATCGTCACATGAATCTTGACCTCCACCAATAAATATTGACAAGC	600						
Db	637	AGGCAAGGACAGGTTATCGTCACATGAATCTTGACCTCCACCAATAAATATTGACAAGC	696						
Qy	601	AGCAATGAAACGACAGAAATGATACGGGATGTTTGGCATCTCCAAAGATACAACTCC	660						
Db	697	AGCAATGAAACGACAGAAATGATACGGGATGTTTGGCATCTCCAAAGATACAACTCC	756						
Qy	661	ACAAACAATCAAAATGCCCCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG	720						
Db	757	ACAAACAATCAAAATGCCCCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG	816						
Qy	721	AGCATTCACCTACAAATCTCAAAATTAATCTGCTAAATCTGGAATCTATGAACCCAGT	780						
Db	817	AGCATTCACCTACAAATCTCAAAATTAATCTGCTAAATCTGGAATCTATGAACCCAGT	876						
Qy	781	AGCGACGATGAGGACCTTATGATTTTCCGGCTCAGGATCTGGAGAACAGGGGCCCAACA	840						
Db	877	AGCGACGATGAGGACCTTATGATTTTCCGGCTCAGGATCTGGAGAACAGGGGCCCAACA	936						
Qy	841	ACTCTTAATGTAGTCACTGAAACAGAAACAAATCGTCAACAATATTTGTCCACTCTTACTA	900						
Db	937	ACTCTTAATGTAGTCACTGAAACAGAAACAAATCGTCAACAATATTTGTCCACTCTTACTA	996						
Qy	901	CATCCAAGCACTTCACAACTAGCAAAACAGTACGAATCTCTCCCGACATGCTGTAACT	960						
Db	997	CATCCAAGCACTTCACAACTAGCAAAACAGTACGAATCTCTCCCGACATGCTGTAACT	1056						
Qy	961	GAGCAAAATGGAAACCGAACCCAAACAACAACAGCAACAGCTCCTCAACAATATCTAATACA	1020						
Db	1057	GAGCAAAATGGAAACCGAACCCAAACAACAACAGCAACAGCTCCTCAACAATATCTAATACA	1116						

RESULT 2

US-08-760-615-3
; Sequence 3, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/POCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2164 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Marburg virus
; STRAIN: Musoke
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 119..2164

US-08-760-615-3
Query Match 74.0%; Score 1514.8; DB 3; Length 2164;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTCTGATTAGTCTCAATTTTAAATCCAAAGTATATAAAATCTCCCT 60
Db 119 ATGAAGACCATATTTCTTCTTATCAGTCTTATCTTAAATTCAGGACACAAAAATCTCCC 178
Qy 61 GTTTTAGAAATGCTAGTAAACAGCCAACTCAAGATGTAGATTGATGCTGCTCCGGAACC 120
Db 179 ATTTTAGAGATAGCTAGTAAATCAACCCAAAATGTGGATTGCGTATGCTCCGGAAC 238
Qy 121 CTCCAAAAGACAGAAGATGTTTCATCTGATGGATTTACACTGAGTGGGCAAAAAGTTGCT 180
Db 239 CTCGAAGACAGAAGAGCTCATCTGATGGATTACATGAGTGGGCAAAAAGTTGCT 298
Qy 181 GATTCCCTTTGGAGCATCTAAACGATGGGCTTTTCCAGACAGATGTTTCTCCCAAGAAC 240
Db 299 GATTCCCTTTGGAGCATCTAAACGATGGGCTTTTCCAGACAGATGTTTCTCCCAAGAAC 358
Qy 241 GTTAGATACGAAGGAGAGAGCCAAACATGTTTCAATATATAGTTTGAACACCT 300

Db 359 GTTGAGTACACAGAGGGGGAGGAGCCAAACATGCTACAATATATAGTTGTAACGGATCCC 418
Qy 301 TCTGAAAAATCTTTGCTGCTGGATCCTCCAGTAAATATCCCGATTAACCTTAATGTAAA 360
Db 419 TCTGAAAAATCTTTGCTGCTGGATCCTCCAGTAAATATCCCGATTAACCTTAATGTAAA 478
Qy 361 ACTGTTTCATCATATATCAAGGTCAAAACCTTCATGACAGAGGGGATGCGCTCATTTGGG 420
Db 479 ACTATCCATCATATATCAAGGTCAAAACCTTCATGACAGAGGGGATGCGCTCATTTGGG 538
Qy 421 GGGGCATTTTCTTGTATGATCGGTTGCTCTCAACAATGTACCGAGGCAAGTCTTTC 480
Db 539 GGAGCATTTTCTTGTATGATCGGTTGCTCTCAACAATGTACCGAGGCAAGTCTTTC 598
Qy 481 ACTGAAGAAATATAGCAGCTATGATTTGTTAATAGACAGTTCAACAGATGATTTTCT 540
Db 599 ACTGAAGGGAACATAGCAGCTATGATTTGTTAATAGACAGTTCAACAGATGATTTTCT 658
Qy 541 AGGCAAGGACAAAGGTTATCGTCATGAACTTGGACCTCCACCAATTAATATTTGGACAGC 600
Db 659 CGGCAAGGACAAAGGTTATCGTCATGAACTTGGACCTCCACCAATTAATATTTGGACAGC 718
Qy 601 AGCAATGAAACGACAGAGAAATGATACGGGATGTTTGGCATCTCTCCAGAAATACAACTCC 660
Db 719 AGTAAACGAGACGCAAAACGAATGACACTGGATGTTTGGCGCTCTTCAAGATACAACTCT 778
Qy 661 ACAAACAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACTCCG 720
Db 779 ACAAAGAACCAAAACATGTTGCTCCGTCCTCAAAATACCTCCACCACTGCCACAGCCGTCG 838
Qy 721 AGCATTCATCTCAAAATACCTCAAAATTAATTAATCTGTAATCTGGAATATGAACTCAAGT 780
Db 839 GAGATCAAACTCAACAGCACCCAACTGATGTCACCAAACTCAATACACGAGCCCAAGC 898
Qy 781 AGGCACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db 899 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGTCGGAGAACAGGAAACCCACACA 958
Qy 841 ACTCTTAATGTAGTCACTGAAACAGAAACAATCGTCAACAATATTTGTCCACTCTTCACTA 900
Db 959 ACTTCTGATGGGTCAACCAAGCAAGGGCTTTTCATCAACAATGCCACCTCCCTCACCA 1018
Qy 901 CATCCAGACCTCAACAACATGAGCAAAAACGATGAGAACTCTTCCGACATCTGTGTAAT 960
Db 1019 CAACCAAGCACGCCACAGCAGAGGAGGAAACAAACAACCAATTCCTCCAGATGCTGTGACT 1078
Qy 961 GAGCACAAATGGAACCGACCCCAACAACAACACCGTCCATGCAACCAATTAATTAACA 1020
Db 1079 GAATAGACAAAATAACAACAACCTGCAACACCGTCCATGCCCCCTCATACACTACCA 1138
Qy 1021 ACTCCACCTATAACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCAAACCCGCAAC 1080
Db 1139 ATCTCTACTAACAACACTCCAAACAACACTTTCAGACTCTCTCTGACCACTTACAAAC 1198
Qy 1081 ATCAACCAATATGATACAAACGTTGAATAGCAGAAAGCGAAACAAACCAATGCTCAGTTG 1140
Db 1199 ACCACCAATGACACACAGAGCAACATCACTGAAATGAGCAAAACAGTGCCTCCCTCG 1258
Qy 1141 AACCAACTCTAGATCCAAAGAAATCCCAACACAGACAGACAGCAACCAACAGCAACACC 1200
Db 1259 ATAACAACCTCTGCTCCAAACGGGAAATCCCAACAGCAAGAGACCAAGAGCAAGCAAAA 1318
Qy 1201 AACATCATCATGACGACATCAGATATAACAGCAAAACACCCCAAAAATTTCTTCCGGAT 1260
Db 1319 GGCCCCGCCCAACGGCACCAACACGCAAAATGAGCATTTCCAGCTCTCTCCCCCACC 1378
Qy 1261 TCTAGTCCGACAAACCCGCTCTTATATATATTTTGAAGAAACGATCGATCTCTGAGG 1320
Db 1379 CCCAGCTCGACTGCACAAACATCTTGTATATTTTCAAGAAAGCGAAGTATCTCTCTGAGG 1438
Qy 1321 GAAGCGCACATGTTCCCTTTCTCGATGGGTTTAAATGCTCCAAATGATTTTGAACCA 1380
Db 1439 GAAGCGCACATGTTCCCTTTCTCGATGGGTTTAAATGCTCCAAATGATTTTGAACCA 1498

```
QY 1381 GTTCAAATACAAAAAATCTTTGATGAATCCTCTAGTCTCTGTCGCTCGGCTGAGGAA 1440
|||
Db 1499 GTTCAAATACAAAAAATCTTTGATGAATCCTCTAGTCTCTGTCGCTCGGCTGAGGAA 1558
|||
QY 1441 GATCAACATGCTCCCTCCCAATATTAGTTTAACTTTATCTTATTTCTTATTTCTTAAATGAG 1500
|||
Db 1559 GATCAACATGCTCCCTCCCAATATTAGTTTAACTTTATCTTATTTCTTATTTCTTAAATGAG 1618
|||
QY 1501 AACACTGCTACTCTGGAGAAAAATGAGAAATGATTTGATGACAGAGCTTAAGAAATTTGGAGC 1560
|||
Db 1619 AACACTGCTACTCTGGAGAAAAATGAGAAATGATTTGATGACAGAGCTTAAGAAATTTGGAGC 1678
|||
QY 1561 GTTCAGGAGATGACTCTGGCCGACGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAATT 1620
|||
Db 1679 GTTCAGGAGATGACTCTGGCCGACGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAATT 1738
|||
QY 1621 GAAGGACTTTACATGCTGTTTTTAATTAATAATCAAAATTTGGCTCTGAGGTTGAGG 1680
|||
Db 1739 GAAGGACTTTACATGCTGTTTTTAATTAATAATCAAAATTTGGCTCTGAGGTTGAGG 1798
|||
QY 1681 CGTCTAGCAATCAAACTGCCAAATCCTTTGGAATCTTTATTGAGAGTCACAACTGAGGAA 1740
|||
Db 1799 CGTCTAGCAATCAAACTGCCAAATCCTTTGGAATCTTTATTGAGAGTCACAACTGAGGAA 1858
|||
QY 1741 AGAACAATTTCTCCTTAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGGAGG 1800
|||
Db 1859 AGAACAATTTCTCCTTAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGGAGG 1918
|||
QY 1801 ACATGCAAAAGTCTTGGACCTGATTGTCATCGGGATAGAGACTTGTCCAAAATATT 1860
|||
Db 1919 ACATGCAAAAGTCTTGGACCTGATTGTCATCGGGATAGAGACTTGTCCAAAATATT 1978
|||
QY 1861 TCAGAGCAAAATGACCAAAATTAAGAGGACGAAACAAAAGAGGGGACTGTTGGGCTG 1920
|||
Db 1979 TCAGAGCAAAATGACCAAAATTAAGAGGACGAAACAAAAGAGGGGACTGTTGGGCTG 2038
|||
QY 1921 GGTGTAATGTTGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTGCTACTA 1980
|||
Db 2039 GGTGTAATGTTGACATCCGACTGGGGTGTCTTACTAACTTGGGCAATTTGCTACTA 2098
|||
QY 1981 TTATCCATAGTCTTGAATGCTCTATCCTGTAATTTGTCGTAATTTTACTAAATATATC 2040
|||
Db 2099 TTATCCATAGTCTTGAATGCTCTATCCTGTAATTTGTCGTAATTTTACTAAATATATC 2158
|||
QY 2041 GGATAA 2046
|||
Db 2159 GGATAA 2164
|||
```

RESULT 3

```
US-09-336-910A-1
; Sequence 1, Application US/09336910A
; Patent No. 6517842
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143
; CURRENT APPLICATION NUMBER: US/09/336,910A
; PRIOR FILING DATE: 1999-06-21
; PRIOR APPLICATION NUMBER: US 60/091,403
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
```

```
; ORGANISM: Marburg Virus
US-09-336-910A-1
```

```
Query Match 74.0%; Score 1514.8; DB 4; Length 11460;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTTCTGATAGTCTCATTTTAAATCCAAAGATATAAAACTCTCCCT 60
Db 5940 ATGAAGACCATATTTTCTTATCAGTCTTATCTTAATTCAGGGACAAAAAATCTCCCC 5999
|||
QY 61 GTTTTAGAATTTCTAGTACAGCCCACTCAGATCTAGATTCAGTGTGCTCCGGAACC 120
|||
Db 6000 ATTTTAGAGATAGCTAGTAATATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 6059
|||
QY 121 CTCCAAAGACAGAAAGATGTTTCATCTGATGGGATTTTACATCTGAGTGGGCAAAAATTTGCT 180
|||
Db 6060 CTCGAAGACAGAAAGAGCTCCATCTGATGGGATTCACATCTGAGTGGGCAAAAATTTGCT 6119
|||
QY 181 GATTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
|||
Db 6120 GATTCCCTTTTGAAGCATCTCAAGCGATGGGCTTTTCAGGACAGGTGACTCCCAAGAAT 6179
|||
QY 241 GTTCAGATACAGGAAGAGAGCCAAACATGTTACATATATAAGTGTAAACAGACCT 300
|||
Db 6180 GTTCAGATACAGAGGGGGAGGAGCCAAACATGCTACAAATATAAGTGTAAACAGATCCC 6239
|||
QY 301 TCTGAAAAATCTTGTCTGCTGGATCTCCCAAGTAAATATCCCGGATTTACCTTAAATGTAAA 360
|||
Db 6240 TCTGAAAAATCTTGTCTGCTGGATCTTACCAACATCCGCTGACTATCTTAATGCAAA 6299
|||
QY 361 ACTGTTTCATATTTCAAGGTCAAAACCTTCATGCACAGGGGATTTGCCCTCCATTTGTGG 420
|||
Db 6300 ACTATCCATCATATTTCAAGGTCAAAACCTTCATGCACAGGGGATTCGCCCTTCATTTATGG 6359
|||
QY 421 GGGGCAATTTTCTGTATGATCGGTTGCCCTCTACAAATGTACCGAGGCAAGGTCTTC 480
|||
Db 6360 GGAGCAATTTTCTGTATGATCGCAATTTGCCCTCCAACAATGTATCCGAGGCAAAATCTTC 6419
|||
QY 481 ACTCAAGCAATATAGCAGCTATGATTTGTTAATAAGACAGTTTACAGAAATGATTTTTTCT 540
|||
Db 6420 ACTCAAGCAATATAGCAGCTATGATTTGTTAATAAGACAGTTTACAGAAATGATTTTTCTG 6479
|||
QY 541 AGCAAGGACAAAGTTATCGTCAATGAATTTGACCTCCACCAATAAATAATTTGGAACAAGC 600
|||
Db 6480 CGGCAAGGACAAAGGTACCGTCAATGAATCTGACTTCTACTAATAAATAATTTGGAACAAGT 6539
|||
QY 601 AGCAATGAACGACAGAAATGATACGGGATGTTTGGCATCTCTCCAGAAATACAACTCC 660
|||
Db 6540 AGTAACGGAAACGCAAAACGAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 6599
|||
QY 661 ACAAACAATCAAAATGCTCCATCTCTTAAACCTCCATCCCTGCCCAAGTAACTCCG 720
|||
Db 6600 ACAAAGAACCAAAATGCTCTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 6659
|||
QY 721 AGCATTTCATCTACAAATATCTCAAAATTAATACTGCTAAATCTGGAACCTATGAACCAAGT 780
|||
Db 6660 GAGATCAAACTCACAAGCACCCCAACTGATGCCCAAACTCAATACCAAGCAAGCAAGC 6719
|||
QY 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAAACAGGGGGCCCCACACA 840
|||
Db 6720 AGTGTATGAGGACCTCGCAACATCTCGGCTCAGGGTCCGAGAAACGAGAAACCCCAACA 6779
|||
QY 841 ACTCTTAATGTAAGTCACTGAACAGAAACAATGCTCAACAATATTGTCACCTCTTCACTA 900
|||
Db 6780 ACTTCTGATGGGTCAACCAAGAGGGCTTTTCATCAACATGCCACCACTCCCTCACC 6839
|||
QY 901 CATCCAGCACCTCACCAACATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAACT 960
|||
Db 6840 CAACCAAGCACGCCACAGCAGGAGGAGAAACAACAAACCACTTCCCAAGATGCTGTGACT 6899
|||
QY 961 GAGCACATGGAACCGACCCCAACACACACAGCAGGCTCTCTCAACAATACTAATACA 1020
|||
```

Db 6900 GAACTAGACAAAATAACACAACTGACACACCGTCCATGCCCCCTCATACACACTACCACA 6959
QY 1021 ACTGCCACCTATACACACTCTCAAGTACAACTCTAGTACTCTTCCCTCCGACCGGAC 1080
Db 6960 ATCTCTACTAACACACCTCCAAACCACTTTCAGCACTCTCTGACCACTTACAAAC 7019
QY 1081 ATACCAATAATGATACACACGCTGAATCTAGCAGAAAGCGAACAAACCAATGCTCAGTTG 1140
Db 7020 ACCACCAATGACACACACAGACCAATCACTGAAATGAGCAACCACTGCCCCCTCG 7079
QY 1141 AACCAACTCTAGATCCAAAGAAAATCCACACAGGACCAAGACACCAACAGCACACAC 1200
Db 7080 ATAACAACCTGCTCCAAAGGAAATCCACACAGCAAGAGAGCACCAGCAGCAAAAA 7139
QY 1201 AACATCATGACGACATCAGATATACAGAGCAACACCCCAAAATCTTCTCCGAT 1260
Db 7140 GGGCCCGCACACGGGACCAACACGCAATGAGCATTTTCCAGCTCTCCGCCACC 7199
QY 1261 TCTAGTCGACACACCGCCCTCTATATACCTTGAAGAAACGATCGATCCTCTGGAGG 1320
Db 7200 CCCAGCTCGATGACAAACATCTTGATATTTTGAAGAAAGCGAAGTATCTCTGGAGG 7259
QY 1321 GAAGCGACATGTTCCCTTTTCTGGATGGGTTAATAATGCTCCAAATGATTTTGACCCA 1380
Db 7260 GAAGCGACATGTTCCCTTTTCTGGATGGGTTAATAATGCTCCAAATGATTTTGACCCA 7319
QY 1381 GTTCCAAATACAAAACAACTTTTGTAGTAATCTCTAGTCTGGTCCCTCGGCTGAGAA 1440
Db 7320 GTTCCAAATACAAAACAACTTTTGTAGTAATCTCTAGTCTGGTCCCTCGGCTGAGAA 7379
QY 1441 GATCAATGCTCCCTCCCAATATAGTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTT 1500
Db 7380 GATCAATGCTCCCTCCCAATATAGTTTAACTTTTAACTTTTAACTTTTAACTTTTAACTTT 7439
QY 1501 AACACTGCTACTCTGGAGAAAATGAGAAATGATTTGATGACAGAGTTTAAAGAAATTTGGAGC 1560
Db 7440 AACACTGCTACTCTGGAGAAAATGAGAAATGATTTGATGACAGAGTTTAAAGAAATTTGGAGC 7499
QY 1561 GTTCAGGAGATGACTCGGCGAGGGCTGATGTTGGATACCGTTTTTTTGGCCCTGGAATT 1620
Db 7500 GTTCAGGAGATGACTCGGCGAGGGCTGATGTTGGATACCGTTTTTTTGGCCCTGGAATT 7559
QY 1621 GAAGGACTTTACATGCTGTTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1680
Db 7560 GAAGGACTTTACATGCTGTTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 7619
QY 1681 CGTCTAGCAATCAAACTGCCAAATCCCTTGGAACTCTTATTTGAGAGTCACAACTGAGAA 1740
Db 7620 CGTCTAGCAATCAAACTGCCAAATCCCTTGGAACTCTTATTTGAGAGTCACAACTGAGAA 7679
QY 1741 AGAATTTCTCTTAATCAATAGACATGCTATTTGATCTTCTTACTCAAGATGGGAGGA 1800
Db 7680 AGAATTTCTCTTAATCAATAGACATGCTATTTGATCTTCTTACTCAAGATGGGAGGA 7739
QY 1801 ACATGCAAGTCTTGGACCTGATTTGATCGGATAGAGACTTGTCCAAAATATT 1860
Db 7740 ACATGCAAGTCTTGGACCTGATTTGATCGGATAGAGACTTGTCCAAAATATT 7799
QY 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAGAGGGGACTGTTTGGGCTCG 1920
Db 7800 TCAGAGCAAAATGACCAAAATTAATAAGGACGAAACAAAAGAGGGGACTGTTTGGGCTCG 7859
QY 1921 GGTGTAATGGTGACATCCGACTGGGGTGTCTTACTAACTTTGGGCAATTTTGTCTACTA 1980
Db 7860 GGTGTAATGGTGACATCCGACTGGGGTGTCTTACTAACTTTGGGCAATTTTGTCTACTA 7919
QY 1981 TTATCCATAGTCTTCTGATGCTCTATCTCTGATTTTGTCTGATCTTTTACTAAATATATC 2040
Db 7920 TTATCCATAGTCTTCTGATGCTCTATCTCTGATTTTGTCTGATCTTTTACTAAATATATC 7979
QY 2041 GGATAA 2046
Db 7980 GGATAA 7985

RESULT 4
US-08-760-615-1
; Sequence 1, Application US/08760615
; Patent No. 6200959
; GENERAL INFORMATION:
; APPLICANT: Haynes, Joel R
; APPLICANT: Schmaljohn, Connie S
; APPLICANT: Fuller, Deborah L
; APPLICANT: Schmaljohn, Alan
; APPLICANT: Jahrling, Peter B
; TITLE OF INVENTION: GENETIC INDUCTION OF ANTI-VIRAL IMMUNE
; TITLE OF INVENTION: RESPONSE AND GENETIC VACCINE FOR FILOVIRUS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08760,615
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 110229.91241
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Ebola virus
; STRAIN: Zaire
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 142..2172
; OTHER INFORMATION: /product= "Glycoprotein"
US-08-760-615-1

Query Match 6.5%; Score 134; DB 3; Length 2172;
Best Local Similarity 60.9%; Pred. No. 1.2e-30;
Matches 218; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 1532 ATTCTGATGACAGCTTAAAGAAATTTGGAGCGTTTCAAGGAGATGACTGGCCGACGGGCTCA 1591
Db 1670 AATGCAACCCCTAATTTACATTAATCTAGTCTGAGTGAAGGTGCTGCAATCGGACTGG 1729
QY 1592 GTTGGATACCGTTTTTTTGGCCCTGGAAATTTGAAGGACTTTACACTGCTGTTTTTAATAAAA 1651
Db 1730 CCTGGATACCATATTTCTGGGCGACGACCGAGGAAATTTACATAGAGGGGCTTAATGCACA 1789
QY 1652 ATCAAAACAATTTGGTCTGACGGTTGAGGCGTCTAGCCAACTCAAACTGCCAAATCCTTGG 1711
Db 1790 ATCAAGATGGTTAATCTGTGGTTGAGACAGCTGGCCCAACGAGACGACTCAAGCTCTTC 1849
QY 1712 AACTCTTTATGAGAGTCACAACTGAGGAAAGAAACATTTCTCTTAACTAATAGACATGCTA 1771
Db 1850 AACTGTTCTCTGAGAGCCACAACTGAGTACGCACTTTTCAATCTCTCAACCGTAAGGCA 1909

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Verry, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480,76-1 (HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5163 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-928-361B-4

Query Match 2.6%; Score 54; DB 3; Length 5163;
Best Local Similarity 44.6%; Pred. No. 2e-05;
Matches 213; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Qy 832 CCCACACAACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAATATTGTCCACT 891
Db 544 CCATACACTAAATGTGTTGGAGTGAACACACACACACACACACAACTACTACTACT 603

Qy 892 CTTCACTACATCCAGACCTCAACATGAGCAAAACAGTACGATCCTTCCGGACAT 951
Db 604 ACTACTAGCAGCAACAAACAAACACGACAAACAACTACTACTACTACTACTACT 663

Qy 952 GCTGTAAGTGAACAACTCTAGATCCACAGAAATCCACACAGAAACGCTCTCAACAAT 1011
Db 664 ACTACTAGCAGCAACAAACAAACACGACAAACAACTACTACTACTACTACTACT 663

Qy 952 GCTGTAAGTGAACAACTCTAGATCCACAGAAATCCACACAGAAACGCTCTCAACAAT 1011
Db 664 ACTACTAGCAGCAACAAACAAACACGACAAACAACTACTACTACTACTACTACT 663

Qy 1012 ACTAATACAACTCCACCTATAACACTCTCAAGTACAACTCTCAAGTACAACTCTTCCCTCCA 1071
Db 724 ACAACAAACAGCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 783

Qy 1072 ACCGCAACATCAACAATAATGATACAAAGTGAATAGCAGAAAGGGAACAAACCAAT 1131
Db 784 ACAACAACTACAAACCACTACTACTACTACTACTACTACTACTACTACTACTACTACT 843

Qy 1132 GCTCAGTTGAACAACTCTAGATCCACAGAAATCCACACAGAAACGCTCTCAACAAT 1191
Db 844 ACACTACAAACCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 903

Qy 1192 AGCAACCAACATCATCATGACGACATCAGATATAACAGCAAAACACCCCAACAATTTCT 1251
Db 904 ACACTACCAAGAAACCAACCAACCACTACTACTACTACTACTACTACTACTACTACT 963

Qy 1252 TCTCCGGATTCTAGTCCGACAAACCGCCCTCTATATATCTTTAGAAAGAAACGATCGA 1309
Db 964 ACTAACAACCAACCAACCAACCAACCACTACTACTACTACTACTACTACTACTACTACT 1021

RESULT 8
US-09-588-995A-4
; Sequence 4, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5

; CURRENT APPLICATION NUMBER: US/09/588,995A
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 08/827,171
; PRIOR FILING DATE: 1997-03-27
; PRIOR APPLICATION NUMBER: 08/928,361
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 08/700,651
; PRIOR FILING DATE: 1996-08-14
; PRIOR APPLICATION NUMBER: 08/415,751
; PRIOR FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 5163
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-09-588-995A-4

Query Match 2.6%; Score 54; DB 4; Length 5163;
Best Local Similarity 44.6%; Pred. No. 2e-05;
Matches 213; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Qy 832 CCCACACAACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAATATTGTCCACT 891
Db 544 CCATACACTAAATGTGTTGGAGTGAACACACACACACACACACAACTACTACTACTACT 603

Qy 892 CTTCACTACATCCAGACCTCAACATGAGCAAAACAGTACGATCCTTCCGGACAT 951
Db 604 ACTACTAGCAGCAACAAACAAACACGACAAACAACTACTACTACTACTACTACTACT 663

Qy 952 GCTGTAAGTGAACAACTCTAGATCCACAGAAATCCACACAGAAACGCTCTCAACAAT 1011
Db 664 ACTACTAGCAGCAACAAACAAACACGACAAACAACTACTACTACTACTACTACTACT 723

Qy 1012 ACTAATACAACTCCACCTATAACACTCTCAAGTACAACTCTCAAGTACAACTCTTCCCTCCA 1071
Db 724 ACAACAAACAGCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 783

Qy 1072 ACCGCAACATCAACAATAATGATACAAAGTGAATAGCAGAAAGGGAACAAACCAAT 1131
Db 784 ACAACAACTACAAACCACTACTACTACTACTACTACTACTACTACTACTACTACTACT 843

Qy 1132 GCTCAGTTGAACAACTCTAGATCCACAGAAATCCACACAGAAACGCTCTCAACAAT 1191
Db 844 ACACTACAAACCACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 903

Qy 1192 AGCAACCAACATCATCATGACGACATCAGATATAACAGCAAAACACCCCAACAATTTCT 1251
Db 904 ACACTACCAAGAAACCAACCAACCACTACTACTACTACTACTACTACTACTACTACT 963

Qy 1252 TCTCCGGATTCTAGTCCGACAAACCGCCCTCTATATATCTTTAGAAAGAAACGATCGA 1309
Db 964 ACTAACAACCAACCAACCAACCAACCACTACTACTACTACTACTACTACTACTACTACT 1021

RESULT 9
US-08-700-651-2
; Sequence 2, Application US/08700651B
; Patent No. 6015882
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: LEECH, JAMES
; APPLICANT: NELSON, RICHARD, C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-4 (HV)
; CURRENT APPLICATION NUMBER: US/08/700,651B
; CURRENT FILING DATE: 1997-08-14
; EARLIER APPLICATION NUMBER: 08/415,751
; EARLIER FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 15


```
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 5318
/ TYPE: DNA
/ ORGANISM: Cryptosporidium parvum
US-08-700-651-2

Query Match      2.6%; Score 54; DB 3; Length 5318;
Best Local Similarity 44.6%; Pred. No. 2.1e-05;
Matches 213; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Qy 832 CCCGACACAACTCTTAATGTAGTCTCACTGAACGAAACAATCGTCAACAATATTGTCCACT 891
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 545 CCATACACTAAATGTGTGGAGTGAACACACAAACAACAACAACAACAACAACAACAACA 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 892 CTTTCACTACATCCAAGCACCTCAACACATGAGCAAAACAGTAGCAATCTTCCGACAT 951
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 605 ACTACTAGACAAACAACAACAACAAGACAAACAACAACAACAACAACAACAACAACA 664
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 952 GGTGTAATGAGCACAATGGAAACCGACCAACAACAACAACAACAACAACAACAACA 1011
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 665 ACTACTAGACAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 724
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1012 ACTAATAAATCCACCTATTAACACTCTCAAGTACAACCTCAGTACTCTTCCGCTCCA 1071
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 725 ACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 784
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1072 ACCCGCAACATCCAAATAATGATACACAACGTGAACCTAGCAGAAAGCGAACAACCAAT 1131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 785 ACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 844
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1132 GTCAGTTGAAACACAACTCTAGATCCAAACAGAAAAATCCACACAGGACAGACACCAAC 1191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 ACAACTACACCAACAACCTAGACCAACCAACAACAACAACAACAACAACAACA 904
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1192 AGCAACAACAACATCATGACGACATCAGATATAACAAGCAAAACCCCAACAATTTCT 1251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 905 ACAACTAGCAAGAAACCAACAACAACAACAACAACAACAACAACAACAACAACA 964
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1252 TCTCCGGATTCTAGTCCGACACCCGCCCTCTATATACTTTAGAAAGAAAGATCGA 1309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 965 ACTACAACCCACAAACAACAACAACAACAACAACAACAACAACAACAACAACA 1022
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-08-928-361B-3
; Sequence 3, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS.
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 60/026,062
```

```
/ FILING DATE: 13-SEP-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Verry, Hana
/ REGISTRATION NUMBER: 30,518
/ REFERENCE/DOCKET NUMBER: 480.76-1(HV)
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-324-1677
/ TELEFAX: 650-324-1678
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5318 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
US-08-928-361B-3

Query Match      2.6%; Score 54; DB 3; Length 5318;
Best Local Similarity 44.6%; Pred. No. 2.1e-05;
Matches 213; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Qy 832 CCCGACACAACTCTTAATGTAGTCACTGAACGAAACAATCGTCAACAATATTGTCCACT 891
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 544 CCATACACTAAATGTGTGGAGTGAACACACAAACAACAACAACAACAACAACAACA 603
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 892 CTTTCACTACATCCAAGCACCTCAACACATGAGCAAAACAGTAGCAATCTTCCGACAT 951
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 604 ACTACTAGCAACAACAACAACAACAAGACAAACAACAACAACAACAACAACAACA 663
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 952 GCTGTAATGAGCACAATGGAAACCGACCAACAACAACAACAACAACAACAACAACA 1011
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 664 ACTACTAGCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 723
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1012 ACTAATAAATCCACCTATTAACACTCTCAAGTACAACCTCAGTACTCTTCCGCTCCA 1071
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 724 ACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 783
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1072 ACCCGCAACATCCAAATAATGATACACAACGTGAACCTAGCAGAAAGCGAACAACCAAT 1131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 784 ACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 843
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1132 GTCAGTTGAAACACAACTCTAGATCCAAACAGAAAAATCCACACAGGACAGACACCAAC 1191
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 844 ACAACTACAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 903
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1192 AGCAACAACAACATCATGACGACATCAGATATAACAAGCAAAACCCCAACAATTTCT 1251
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 904 ACAACTACCAAGAAACCAACAACAACAACAACAACAACAACAACAACAACAACA 963
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 1252 TCTCCGGATTCTAGTCCGACAAACCCGCCCTCTATATACTTTAGAAAGAAAGATCGA 1309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 964 ACTACAACCACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1021
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
US-09-588-995A-3
; Sequence 3, Application US/09588995A
; Patent No. 6514697
; GENERAL INFORMATION:
; APPLICANT: PETERSEN, CAROLYN
; APPLICANT: BARNES, DEBRA A.
; APPLICANT: NELSON, RICHARD C.
; APPLICANT: GUT, JIRI
; TITLE OF INVENTION: METHODS FOR DETECTION OF CRYPTOSPORIDIUM SPECIES AND
; TITLE OF INVENTION: ISOLATES AND FOR DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: INFECTIONS
; FILE REFERENCE: 480.19-5
; CURRENT APPLICATION NUMBER: US/09/588,995A
; CURRENT FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 08/827,171
; PRIOR FILING DATE: 1997-03-27
; PRIOR APPLICATION NUMBER: 08/928,361
; PRIOR FILING DATE: 1997-09-12
```

```
; PRIOR APPLICATION NUMBER: 08/700,651
; PRIOR FILING DATE: 1996-08-14
; PRIOR APPLICATION NUMBER: 08/415,751
; PRIOR FILING DATE: 1995-04-03
; NUMBER OF SEQ ID NOS: 115
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5318
; TYPE: DNA
; ORGANISM: Cryptosporidium parvum
US-09-588-995A-3

Query Match      2.4%; Score 54; DB 4; Length 5318;
Best Local Similarity 44.6%; Pred. No. 2.1e-05;
Matches 213; Conservative 0; Mismatches 265; Indels 0; Gaps 0;

Qy 832 CCCACACAACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAAATATTTGCCACT 891
Db 544 CCATACACTAAATGTGTGGAGTGAACACACAAACAACAACAACAACAACAACAACAACA 603

Qy 892 CTTTCACTACATCCAGACCTCTCAACATGAGCAAAACAGTACGATCTCTCCGGACAT 951
Db 604 ACTACTAGCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 663

Qy 952 GCTGTAACTGAGCACAATGGAAACCGACCCCAACAACAACAACAACAACAACAACAACA 1011
Db 664 ACTACTAGCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 723

Qy 1012 ACTAATACAACTCCACCTATAAACAATCTCAAGTACAACTCAGTACTCTCTTCCCTCCA 1071
Db 724 ACAACAAGACGACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACC 783

Qy 1072 ACCGGCAATCACAATAATGATATACAAACGTAAGTAACTAGCAGAAAGGGAACAACCAAT 1131
Db 784 ACAACAATCACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 843

Qy 1132 GCTCAGTTGAACACAACCTCTAGATCCACAGAAATCCACACAGGACAGACCAAC 1191
Db 844 ACAACTACACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 903

Qy 1192 AGCAACAACAACATCATCATGACGACATCAGATATAACAAGCAACACCCCAAAATCT 1251
Db 904 ACAACTACAGAAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAAC 963

Qy 1252 TCTCGGATTCAGTCCGACAAACCGCCCTCCTATATATCTTTAGAAAGAAACGATCGA 1309
Db 964 ACTACAAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1021

RESULT 12
US-09-270-767-2848/c
; Sequence 2848, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2848
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: n means any nucleotide
US-09-270-767-2848

Query Match      2.4%; Score 48.8; DB 4; Length 631;
Best Local Similarity 51.1%; Pred. No. 0.0002;
Matches 139; Conservative 0; Mismatches 132; Indels 1; Gaps 1;

Qy 977 ACCCAACAACAACACACACACGCTCTCAACATATCTAATACAACTCCCACTATACA 1036
Db 628 AACCAACAACAACACACACACGCTCTCAACATATCTAATACAACTCCCACTATACA 569

Qy 1037 CTCTCAAGTACAACTCCTAGTACTCTCTCCCTCCAAACCGCAACATCAACCAATATGATA 1096
Db 568 AACCAACAACAACACACACACGCTCTCAACATATCTAATACAACTCCCACTATACA 509

Qy 1097 CACAACGTGACTAGCAGAAAGCGA-ACAAACCAATGCTCAGTTGAAACAGAACTCTAGAT 1155
Db 508 AAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 449

Qy 1156 CCAACAGAAATCCCAACAGGACAAAGACACCAACAGCAGCAACCAACATCATCATGACG 1215
Db 448 CAAACAACATACCAACTCAAAAAACCAACAACCAACCAACCAACCAACCAACCAACCC 389

Qy 1216 ACATCAGATATAAACAAGCAACACCCCAACAA 1247
Db 388 AACCAAACTACAAACAACAACAACAACAACAACA 357

RESULT 13
US-09-270-767-18130/c
; Sequence 18130, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18130
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: n means any nucleotide
US-09-270-767-18130

Query Match      2.4%; Score 48.8; DB 4; Length 631;
Best Local Similarity 51.1%; Pred. No. 0.0002;
Matches 139; Conservative 0; Mismatches 132; Indels 1; Gaps 1;

Qy 977 ACCCAACAACAACACACACGCTCTCAACATATCTAATACAACTCCCACTATACA 1036
Db 628 AACCAACAACAACACACACGCTCTCAACATATCTAATACAACTCCCACTATACA 569

Qy 1037 CTCTCAAGTACAACTCCTAGTACTCTCTCCCTCCAAACCGCAACATCAACCAATATGATA 1096
Db 568 AACCAACAACAACACACACGCTCTCAACATATCTAATACAACTCCCACTATACA 509

Qy 1097 CACAACGTGACTAGCAGAAAGCGA-ACAAACCAATGCTCAGTTGAAACAGAACTCTAGAT 1155
Db 508 AAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 449

Qy 1156 CCAACAGAAATCCCAACAGGACAAAGACACCAACAGCAGCAACCAACATCATCATGACG 1215
Db 448 CAAACAACATACCAACTCAAAAAACCAACAACCAACCAACCAACCAACCAACCAACCC 389

Qy 1216 ACATCAGATATAAACAAGCAACACCCCAACAA 1247
Db 388 AACCAAACTACAAACAACAACAACAACAACAACA 357

RESULT 14
US-08-928-361B-2
; Sequence 2, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
```

```

; TITLE OF INVENTION:  THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
;
; TITLE OF INVENTION:  FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
;
; TITLE OF INVENTION:  SPECIES INFECTIONS
;
; NUMBER OF SEQUENCES:  30
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE:  PETERS, VERNY, JONES & BIKSA
;
; STREET:  385 Sherman Avenue, Suite 6
;
; CITY:  Palo Alto
;
; STATE:  CA
;
; COUNTRY:  USA
;
; ZIP:  94306-1840
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE:  Floppy disk
;
; COMPUTER:  IBM PC compatible
;
; OPERATING SYSTEM:  PC-DOS/MS-DOS
;
; SOFTWARE:  PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER:  US/08/928,361B
;
; FILING DATE:  12-SEP-1997
;
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
;
; APPLICATION NUMBER:  US 60/026,062
;
; FILING DATE:  13-SEP-1996
;
; ATTORNEY/AGENT INFORMATION:
;
; NAME:  Verny, Hana
;
; REGISTRATION NUMBER:  30,518
;
; REFERENCE/DOCKET NUMBER:  480.76-1 (HV)
;
; TELECOMMUNICATION INFORMATION:
;
; TELEPHONE:  650-324-1677
;
; TELEFAX:  650-324-1678
;
; INFORMATION FOR SEQ ID NO:  2:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH:  5511 base pairs
;
; TYPE:  nucleic acid
;
; STRANDEDNESS:  double
;
; TOPOLOGY:  linear
;
; MOLECULE TYPE:  DNA (genomic)
;
; US-08-928-361B-2

```

Query Match	2.3%	Score 48	DB 3	Length 5511
Best Local Similarity	44.2%	Pred. No. 0.0016		
Matches 198	Conservative 0	Mismatches 250	Indels 0	Gaps 0
Qy	838	ACAACCTCTTAATGTFAGTGCAC	TGACAGAAACAATCGTCAACA	ATATTTGTCACCTCCCTTCA 897
Db	988	ACAACAACCTACTACACTACTCT	ACACCTACTACGACACACACT	ACTACTACTACTACTACGACA 1047
Qy	898	CTACATCCAAAGCACCTCACA	CAACATGAGCAAAAACAGTAC	CGAATCCTTCCGACATGCTGTA 957
Db	1048	ACAACAACAACAACTACTACT	TACTACTACAACGACAAACA	ACTACAACGACAACTTACA 1107
Qy	958	ACTGAGCACAAATGGAAACG	AGCCCAACAACAACAACG	ACAGGTCCTCAACAATACTAAT 1017
Db	1108	ACCAACAACCAACTACTCCA	AGAAACCAACAACAACA	CAACAACAACACTACTACTACT 1167
Qy	1018	ACAACCTCCACCTATAAC	ACTCTCAAGTACAACTCT	CCTAGTACTCTCTCCCTCCAAACCCGC 1077
Db	1168	ACAACAACAACAACAACA	CAACTACTACTACTACT	ACTACTACTACTACTACTACT 1227
Qy	1078	AACATCAACCAATATATG	ATACACAACGTGTAAC	TAGCAGAAAACGCAACAAACCAATGTCTAG 1137
Db	1228	ACTACCAACAACAACCA	CAACCAACCAACTAC	CAAGAAACCAACAACAACAACA 1287
Qy	1138	TTGAACAACAACCTTAG	ATGCGAAGAGAAATCC	ACACGAGGACACGACACGACACA 1197
Db	1288	ACAACAACAACAACCT	TACTACTACAACACG	ACACACAACAACCAACCAACCA 1347
Qy	1198	ACCAACATCATCATG	ACGACATCAGATAT	ATACAGAGCAACACCCCAAAATTTCTTCCG 1257
Db	1348	ACTACAACCTACCA	AGAAACCAACCTACT	ACTACTACTACTACTACCAACAACAACACTACT 1407
Qy	1258	GATTCTAGTCCGACA	ACCCGCGCTCTCTTA	1285

Search completed: August 1, 2005, 10:31:17
Job time : 332.71 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 16:57:35 ; Search time 1285.84 Seconds
(without alignments)
10298.647 Million cell updates/sec

Title: US-10-066-506A-7

Perfect score: 2046

Sequence: 1 atgaagaccatatatttct.....ttactaatatcggataa 2046

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2046	100.0	2046	15 US-10-066-506A-7	Sequence 7, Appl
2	1847.6	90.3	2046	15 US-10-066-506A-13	Sequence 13, Appl
3	1514.8	74.0	7778	20 US-10-491-121-30	Sequence 30, Appl
4	1514.8	74.0	11460	15 US-10-267-322-1	Sequence 1, Appl
5	1485.6	72.6	2051	15 US-10-066-506A-9	Sequence 9, Appl
6	1473.2	72.0	2046	21 US-10-811-353-2	Sequence 2, Appl
7	1473.2	72.0	19112	17 US-10-353-856-27	Sequence 27, Appl

8	1418.8	69.3	7005	20 US-10-491-121-31	Sequence 31, Appl
9	1418.8	69.3	8256	20 US-10-491-121-32	Sequence 32, Appl
10	1319.6	64.5	2046	15 US-10-066-506A-5	Sequence 5, Appl
11	834	40.8	6902	20 US-10-491-121-42	Sequence 42, Appl
12	780	38.1	1841	15 US-10-066-506A-3	Sequence 3, Appl
13	746.6	36.5	2252	15 US-10-066-506A-1	Sequence 1, Appl
14	149	7.3	18890	17 US-10-353-856-9	Sequence 9, Appl
15	148.2	7.2	6885	20 US-10-491-121-18	Sequence 18, Appl
16	148.2	7.2	6889	20 US-10-491-121-19	Sequence 19, Appl
17	148.2	7.2	7002	20 US-10-491-121-16	Sequence 16, Appl
18	148.2	7.2	7003	22 US-10-860-878-1	Sequence 1, Appl
19	148.2	7.2	7023	20 US-10-491-121-21	Sequence 21, Appl
20	148.2	7.2	7036	20 US-10-491-121-17	Sequence 17, Appl
21	148.2	7.2	8146	20 US-10-491-121-20	Sequence 20, Appl
22	145.8	7.1	2360	17 US-10-397-635-12	Sequence 12, Appl
23	145.8	7.1	6913	20 US-10-491-121-11	Sequence 11, Appl
24	145.8	7.1	8131	20 US-10-491-121-12	Sequence 12, Appl
25	134.4	6.6	6940	20 US-10-491-121-15	Sequence 15, Appl
26	134.4	6.6	7073	22 US-10-860-878-2	Sequence 2, Appl
27	134.4	6.6	7082	20 US-10-491-121-13	Sequence 13, Appl
28	134.4	6.6	7087	20 US-10-491-121-14	Sequence 14, Appl
29	134.4	6.6	8221	20 US-10-491-121-27	Sequence 27, Appl
30	134.4	6.6	8338	20 US-10-491-121-26	Sequence 26, Appl
31	134.4	6.6	10783	20 US-10-491-121-25	Sequence 25, Appl
32	134	6.5	2039	15 US-10-066-506A-11	Sequence 11, Appl
33	134	6.5	2298	9 US-09-337-946A-1	Sequence 1, Appl
34	134	6.5	2298	17 US-10-384-976-1	Sequence 1, Appl
35	134	6.5	2298	18 US-10-226-795-1	Sequence 1, Appl
36	134	6.5	2298	22 US-10-696-633-1	Sequence 1, Appl
37	134	6.5	6467	20 US-10-491-121-10	Sequence 10, Appl
38	134	6.5	6561	20 US-10-491-121-4	Sequence 4, Appl
39	134	6.5	6624	20 US-10-491-121-3	Sequence 3, Appl
40	134	6.5	6914	20 US-10-491-121-9	Sequence 9, Appl
41	134	6.5	7154	20 US-10-491-121-1	Sequence 1, Appl
42	134	6.5	7188	20 US-10-491-121-2	Sequence 2, Appl
43	134	6.5	7272	22 US-10-860-878-4	Sequence 4, Appl
44	134	6.5	7285	22 US-10-860-878-3	Sequence 3, Appl
45	134	6.5	8199	20 US-10-491-121-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-10-066-506A-7
; Sequence 7, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION: Case C.
; APPLICANT: Grogan, Michael C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ IDS NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 7
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Raven Glycoprotein
; OTHER INFORMATION: Marburg virus strain Musoko Glycoprotein 2
US-10-066-506A-7

Query Match 100.0%; Score 2046; DB 15; Length 2046;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 60

Db 1 ATGAAGACCATATATTTCTGATGATCTCAATTTTAATCCAAAGTATATAAAACTCTCCCT 60
Qy 61 GTTTTATAGAAATTTGCTAGTAAACAGCCCAACTCAAGATGTAGATTCAGTGTCTCCGGAAAC 120
Db 61 GTTTTATAGAAATTTGCTAGTAAACAGCCCAACTCAAGATGTAGATTCAGTGTCTCCGGAAAC 120
Qy 121 CTCCTAAAGACAGAAAGATGTTTCATCTGATGGATTTTACATGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCCTAAAGACAGAAAGATGTTTCATCTGATGGATTTTACATGAGTGGGCAAAAAGTTGCT 180
Qy 181 GATTCCCTTTTGGAGCATCTAAACGATGGGCTTTTCCAGACAGTGTCTCCCAAGAAC 240
Db 181 GATTCCCTTTTGGAGCATCTAAACGATGGGCTTTTCCAGACAGTGTCTCCCAAGAAC 240
Qy 241 GTTGTAGTATACGGAAGGAGAGAGCCAAACATGTTACAATATAAGTGTAAACAGACCT 300
Db 241 GTTGTAGTATACGGAAGGAGAGAGCCAAACATGTTACAATATAAGTGTAAACAGACCT 300
Qy 301 TCTGGAATCCTTTGCTGCTGATCCTCCAGTAAATATCCGCAATACCTTAAATGTAAA 360
Db 301 TCTGGAATCCTTTGCTGCTGATCCTCCAGTAAATATCCGCAATACCTTAAATGTAAA 360
Qy 361 ACTGTTTATATATCAAGGTCAAAACCTCATGACAGGGGATTCCTCATTTTGG 420
Db 361 ACTGTTTATATATCAAGGTCAAAACCTCATGACAGGGGATTCCTCATTTTGG 420
Qy 421 GGGGCAATTTTCTGTATGATGCGTTCCTCTACAACAATGTACCGAGGCAAGTCTTC 480
Db 421 GGGGCAATTTTCTGTATGATGCGTTCCTCTACAACAATGTACCGAGGCAAGTCTTC 480
Qy 481 ACTGAAGGAAATATATAGCAGCTATGATGTTTAAATAGACAGTTTCAAGAAATGATTTTCT 540
Db 481 ACTGAAGGAAATATATAGCAGCTATGATGTTTAAATAGACAGTTTCAAGAAATGATTTTCT 540
Qy 541 AGGCAAGGACAGGTTATCGTCAATGAACTTTGACCTCCACCAATATAATATTTGACAAAGC 600
Db 541 AGGCAAGGACAGGTTATCGTCAATGAACTTTGACCTCCACCAATATAATATTTGACAAAGC 600
Qy 601 AGCAATGAACGACAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAATCTCC 660
Db 601 AGCAATGAACGACAGAAATGATACGGGATGTTTGGCATCTCCCAAGAAATACAATCTCC 660
Qy 661 ACAACAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAATCCG 720
Db 661 ACAACAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAATCCG 720
Qy 721 AGCATTCTCTACAATACTCAAAATTAATGCTTAATCTGTAATCTGGAACCTATGAACCCAAGT 780
Db 721 AGCATTCTCTACAATACTCAAAATTAATGCTTAATCTGTAATCTGGAACCTATGAACCCAAGT 780
Qy 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAACAGGGGCCCCACACA 840
Db 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAACAGGGGCCCCACACA 840
Qy 841 ACTCTTAATGTAGTCACTGAAACAGAAACAATCGTCAACAATATTTGTCACCTCTTCACTA 900
Db 841 ACTCTTAATGTAGTCACTGAAACAGAAACAATCGTCAACAATATTTGTCACCTCTTCACTA 900
Qy 901 CATCAAGCACTCAACAATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAACT 960
Db 901 CATCAAGCACTCAACAATGAGCAAAACAGTACGAATCTTCCCGACATGCTGTAACT 960
Qy 961 GAGCAATATGGAACCGCCACACACACACAGCAACGCTCTCAACAANTACTAATACA 1020
Db 961 GAGCAATATGGAACCGCCACACACACACAGCAACGCTCTCAACAANTACTAATACA 1020
Qy 1021 ACTCCACCTATAACACTCTCAAGTACAACTCTCAGTACTCTTCCCTCCCAACCGCAAC 1080
Db 1021 ACTCCACCTATAACACTCTCAAGTACAACTCTCAGTACTCTTCCCTCCCAACCGCAAC 1080
Qy 1081 ATCACAATATATGATACACAAGTGAATAGCAGAAAGGGAACCAACCAATGCTCAGTTG 1140

Db 1081 ATCACAATATATGATACACAAGTGAATAGCAGAAAGCGAACAAACCAATGCTCAGTTG 1140
Qy 1141 AACACAACCTCTAGATCCAAACAGAAAAATCCCAACAGGACAAAGACACCAACAGACACAC 1200
Db 1141 AACACAACCTCTAGATCCAAACAGAAAAATCCCAACAGGACAAAGACACCAACAGACACAC 1200
Qy 1201 AACATCATGATGAGGACATCAGATATATAAAGAAAGAAACACCCCAAAATTTCTTCCGGAT 1260
Db 1201 AACATCATGATGAGGACATCAGATATATAAAGAAAGAAACACCCCAAAATTTCTTCCGGAT 1260
Qy 1261 TCTAGTCCGACAAACCCCTCTATATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
Db 1261 TCTAGTCCGACAAACCCCTCTATATATCTTTAGAAAGAAACGATCGATCTCTGGAGG 1320
Qy 1321 GAAGGCGACATGTTTCCCTTTTCTGGATGGGTTAAATAATGCTCCAAATGATTTTGAACCA 1380
Db 1321 GAAGGCGACATGTTTCCCTTTTCTGGATGGGTTAAATAATGCTCCAAATGATTTTGAACCA 1380
Qy 1381 GTTCCAAATACAAAAACAATCTTTGATGAATCCCTAGTTCGTGCTCGGCTGAGGAA 1440
Db 1381 GTTCCAAATACAAAAACAATCTTTGATGAATCCCTAGTTCGTGCTCGGCTGAGGAA 1440
Qy 1441 GATCAACATGCTCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTAATATAAATGAG 1500
Db 1441 GATCAACATGCTCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTAATATAAATGAG 1500
Qy 1501 AACACTGCTACTCTCGGAGAAATGAGATGATTTGATGACAGAGTTAAGAAATTTGGAGC 1560
Db 1501 AACACTGCTACTCTCGGAGAAATGAGATGATTTGATGACAGAGTTAAGAAATTTGGAGC 1560
Qy 1561 GTTCAGGAGGATGACCTGGCCGACAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAT 1620
Db 1561 GTTCAGGAGGATGACCTGGCCGACAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAT 1620
Qy 1621 GAAGGACTTTTACACTGCTGTTTTTAAATTAATAAATCAAAACAATTTGGTCTCGAGTTGAGG 1680
Db 1621 GAAGGACTTTTACACTGCTGTTTTTAAATTAATAAATCAAAACAATTTGGTCTCGAGTTGAGG 1680
Qy 1681 CGTCTAGCCCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCAACAATGAGGAA 1740
Db 1681 CGTCTAGCCCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGAGTCAACAATGAGGAA 1740
Qy 1741 AGAACATTTCTTTAATCAATAGACATGCTTATGACTTTTACTCAAGAATGGGAGGA 1800
Db 1741 AGAACATTTCTTTAATCAATAGACATGCTTATGACTTTTACTCAAGAATGGGAGGA 1800
Qy 1801 ACATGCAAGTGTCTGGACCTGATTGTCATCGGATAGAGACTTCTCCAAAAATATT 1860
Db 1801 ACATGCAAGTGTCTGGACCTGATTGTCATCGGATAGAGACTTGTCCAAAAATATT 1860
Qy 1861 TCAGAGCAAAATTTGACCAAAATTTAAAAAGGACGAAACAAAAAGAGGGGACTGGTTGGGCTG 1920
Db 1861 TCAGAGCAAAATTTGACCAAAATTTAAAAAGGACGAAACAAAAAGAGGGGACTGGTTGGGCTG 1920
Qy 1921 GGTGGTAAATGGTGGACATCCGACTGGGGTCTTACTAATCTTGGGCAATTTGCTACTA 1980
Db 1921 GGTGGTAAATGGTGGACATCCGACTGGGGTCTTACTAATCTTGGGCAATTTGCTACTA 1980
Qy 1981 TTATCCATAGTGTCTTGATGCTCTATCCTGTATTTGTCGTATCTTTTACTAAATATATC 2040
Db 1981 TTATCCATAGTGTCTTGATGCTCTATCCTGTATTTGTCGTATCTTTTACTAAATATATC 2040
Qy 2041 GGATAA 2046
Db 2041 GGATAA 2046

RESULT 2
US-10-066-506A-13
; Sequence 13, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.

APPLICANT: Hevey, Michael C.
Schmaljohn, Alan, L.
TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
FILE REFERENCE: 003/243/SAP
CURRENT APPLICATION NUMBER: US/10/066,506A
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/267,522
PRIOR FILING DATE: 2001-01-31
NUMBER OF SEQ ID NOS: 30
SOFTWARE: Apple Macintosh Microsoft Word 6.0
SEQ ID NO 13
LENGTH: 2046
TYPE: DNA
ORGANISM: Marburg virus strain Raven

OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 and Marburg

OTHER INFORMATION: Glycoprotein 2

US-10-066-506A-13

Query Match 90.3%; Score 1847.6; DB 15; Length 2046;

Best Local Similarity 93.9%; Pred. No. 0;

Matches 1922; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

Qy	1	ATGAAGACCATATTTCTGATTAGTCTCAATTTTAAATCCAAAGTATAAAATCTCCCT	60
Db	1	ATGAAGACCATATTTCTGATTAGTCTCAATTTTAAATCCAAAGTATAAAATCTCCCT	60
Qy	61	GTTTTAGAAATGTGTAGTAACAGCAACCTCAAGATGTAGATTCAGTGTGTCGCGAAC	120
Db	61	GTTTTAGAAATGTGTAGTAACAGCAACCTCAAGATGTAGATTCAGTGTGTCGCGAAC	120
Qy	121	CTCCAAAGACAGAGATGTTCTGATGAGGATTTACATGAGTGGGCAAAAGTTGCT	180
Db	121	CTCCAAAGACAGAGATGTTCTGATGAGGATTTACATGAGTGGGCAAAAGTTGCT	180
Qy	181	GATTCCTCTTGGAGCATCTAAGCATGGGCTTTCAGGACAGGTGTTCTCCCAAGAC	240
Db	181	GATTCCTCTTGGAGCATCTAAGCATGGGCTTTCAGGACAGGTGTTCTCCCAAGAC	240
Qy	241	GTTCAGTATACGGAAGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT	300
Db	241	GTTCAGTATACGGAAGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT	300
Qy	301	TCTGGAAATCTTGCTGCTGATCTCCAGTAAATATCCGGAATACCTTAATGTAAA	360
Db	301	TCTGGAAATCTTGCTGCTGATCTCCAGTAAATATCCGGAATACCTTAATGTAAA	360
Qy	361	ACTGTTCAATATTCAGGTCRAAACCTCATGCACAGGGGATTCCTCCATTTGTTGG	420
Db	361	ACTGTTCAATATTCAGGTCRAAACCTCATGCACAGGGGATTCCTCCATTTGTTGG	420
Qy	421	GGGGCAATTTTCTGTATGATCGGCTTGCCTCTACCAATATGTAACCGAGCAAGTCTTC	480
Db	421	GGGGCAATTTTCTGTATGATCGGCTTGCCTCTACCAATATGTAACCGAGCAAGTCTTC	480
Qy	481	ACTGAAGAAATATAGCAGCTATGATGTTTAAATAGACAGTTCAAGAAATGATTTTCT	540
Db	481	ACTGAAGAAATATAGCAGCTATGATGTTTAAATAGACAGTTCAAGAAATGATTTTCT	540
Qy	541	AGGCAAGACAGGTTATCGTCACATGAACTTGACCTCCACCAATATATGGACAGC	600
Db	541	AGGCAAGACAGGTTATCGTCACATGAACTTGACCTCCACCAATATATGGACAGC	600
Qy	601	AGCAATGAACCGCAGAGAAATGATACGGATGTTTGGCATCTCCCAAGATACAACTCC	660
Db	601	AGCAATGAACCGCAGAGAAATGATACGGATGTTTGGCATCTCCCAAGATACAACTCC	660
Qy	661	ACAAACAATCAACATGCCCTCATCTCTTAACCTCCATCCCTGCCACAGTAATCCG	720
Db	661	ACAAACAATCAACATGCCCTCATCTCTTAACCTCCATCCCTGCCACAGTAATCCG	720
Qy	721	AGCATTCATCTACAATATCTCAATTAATTAATCTGCTAAATCTGGAACATGAAACCAAGT	780

Db	721	AGCATTCATCTACAAATACTCAATTAATTAATCTGCTAAATCTGGAACATATGAACCCCAAGT	780
Qy	781	AGGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA	840
Db	781	AGGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA	840
Qy	841	ACTCTTAATCTAGTCTACGACGAGAACAACTCGTCAACAATATTTGTCCACTCTCTCACTA	900
Db	841	ACTCTTAATCTAGTCTACGACGAGAACAACTCGTCAACAATATTTGTCCACTCTCTCACTA	900
Qy	901	CATCCAGCACCTTCAACAACATGAGCAAAACAGTACGAATCTCTCCGACATGCTGTAACT	960
Db	901	CATCCAGCACCTTCAACAACATGAGCAAAACAGTACGAATCTCTCCGACATGCTGTAACT	960
Qy	961	GAGCACAATGGAACCGACCCCAACAACAGTACGAATCTCTCCGACATGCTGTAACT	1020
Db	961	GAGCACAATGGAACCGACCCCAACAACAGTACGAATCTCTCCGACATGCTGTAACT	1020
Qy	1021	ACTCCACCTTATACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCCAACCCGCAAC	1080
Db	1021	ACTCCACCTTATACACTCTCAAGTACAACTCAGTACTCTCTCCCTCCCAACCCGCAAC	1080
Qy	1081	ATCACCAATATGATACACAACGTTGAATAGCAGAAAGCGAAACAAACCAATGCTCAGTTG	1140
Db	1081	ATCACCAATATGATACACAACGTTGAATAGCAGAAAGCGAAACAAACCAATGCTCAGTTG	1140
Qy	1141	AACACAATCTAGATCCAAAGAAAATCCCAACAAGGACAGACCAACAGCACAAC	1200
Db	1141	AACACAATCTAGATCCAAAGAAAATCCCAACAAGGACAGACCAACAGCACAAC	1200
Qy	1201	AACATCATGACGACATCAGATATACAGCAAAACACCCCAACAAATCTCTCCGAT	1260
Db	1201	AACATCATGACGACATCAGATATACAGCAAAACACCCCAACAAATCTCTCCGAT	1260
Qy	1261	TCTAGTCCGACAAACCCGCTCTCTATATATCTTGAAGAAAGAACGATCGATCCTCTGAGG	1320
Db	1261	TCTAGTCCGACAAACCCGCTCTCTATATATCTTGAAGAAAGAACGATCGATCCTCTGAGG	1320
Qy	1321	GAAGGACATGTTCCCTTTCTGGATGGGTTAATAATGCTCCAAATGATTTTGACCCA	1380
Db	1321	GAAGGACATGTTCCCTTTCTGGATGGGTTAATAATACTGAATGATTTTGATGATCA	1380
Qy	1381	GTTCCAATACAAAACAACTCTTGAATGATCTCTAGTTCTGGTCTCGGCTGAGAA	1440
Db	1381	ATCCCAACACAGAAACAACTCTTGAATGATCTCTAGTTCTGGTCTCGGCTGAGAA	1440
Qy	1441	GATCAACATGCTCTCCCAATATGATTTAACTTTATCTTATTTTCTTAATAAATGAG	1500
Db	1441	GAACACACACTCCCGCATATCAGTTTAACTTTCTTATTTCTGATATAAATGGA	1500
Qy	1501	AACACTGCTACTCTGGAGAAATGAGATGATGATGATGATGATGATGATGATGATGATG	1560
Db	1501	GATACTGCTACTCTGGAGAAACGAGATGATGATGATGATGATGATGATGATGATGATG	1560
Qy	1561	GTTCAGAGGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTGGGCTGGGATTT	1620
Db	1561	GTTCAGAGGATGACCTGGCGGAGGCTCAGTTGGATACCGTTTGGGCTGGGATTT	1620
Qy	1621	GAAGGACTTTACCTGCTGTTTAAATTAATAAATCAAAACAAATTTGGTCTGAGGTTGAGG	1680
Db	1621	GAAGGACTTTACCTGCTGTTTAAATTAATAAATCAAAACAAATTTAGTTGTTGTTGAGG	1680
Qy	1681	CGTCTAGCAATCAAACTGCCAAATCTCTTGAATCTCTTATTTGAGAGTCAACCTGAGAA	1740
Db	1681	CGCTTAGCAATCAAACTGCCAAATCTCTTGAATCTCTTATTTGAGAGTCAACCTGAGAA	1740
Qy	1741	AGAACAATCTCTTAATCAATAGACATGCTATGATCTCTTCTACTCACAAGATGGGGAGA	1800
Db	1741	AGGACATTTCTTAATCAATAGGATGATGATGATGATGATGATGATGATGATGATGATG	1800
Qy	1801	ACATGCAAGGCTTGGACCTGATTTGTCATCGGATAGAGACTTCTGTCACAAATATTT	1860
Db	1801	ACATGCAAGGCTTGGACCTGATTTGTCATCGGATAGAGACTTCTATCTATAAATATCT	1860

```
QY 1861 TCAGAGCAAAATTGACCAAAATTAAGGACGAAACAAAGAGGGGACTGGTTGGGGTCTG 1920
DB 1861 TCAGAACAAATCGACAAATCAGAAAGGATGACAAAGAGGAGAACTGGCTGGGGTCTA 1920
QY 1921 GGTGTAATGGTGACATCCGACTGGGGTGTCTTAACTAACTGGGGCAATTTGGCTACTA 1980
DB 1921 GGTGCAATGGTGACATCTGACTGGGGTGTCTTCACTAAATTTGGGCATCTCTGCTACTA 1980
QY 1981 TTATCCATAGCTGCTTCAATGCTCTATCCGTATTTCTGCTATCTTTTACTAAATATATC 2040
DB 1981 TTATCTATAGCTGTTCTGATGCTCTGCTGCTATCTGCTGATCTTCTCACTAAATACATT 2040
QY 2041 GGATAA 2046
DB 2041 GGATGA 2046

RESULT 3
US-10-491-121-30
; Sequence 30, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 7778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012 Marburg
US-10-491-121-30

Query Match 74.0%; Score 1514.8; DB 20; Length 7778;
Best Local Similarity 83.8%; Pred. No. 0;
Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAACTCTCCCT 60
DB 2034 ATGAAGACCATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAACTCTCCCT 2093
QY 61 GTTTTAAAGATTTGCTAGTAACCAACCTCAAGATGTAGATTTCAGTGTGCTCCGGAAACC 120
DB 2094 ATTTTAAAGATAGCTAGTAATAATCAACCCCAATATGGATTGGTATGCTCCGGAACT 2153
QY 121 CTCGAAAGACGAAGATGTTTCATCTGATGGATTTTACATGAGTGGGCAAAAAGTTGCT 180
DB 2154 CTCGAAAGACGAAGACGTCATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT 2213
QY 181 GATTCCTCTTTGGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
DB 2214 GATTCCTCTTTGGAAGCATCTAAACGATGGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 2273
QY 241 GTTGTAGTATACGGAAGGAGAAAGCAACCAATGTTTACATATATAGTGTAAACGACCT 300
DB 2274 GTTGTAGTATACGGAAGGAGGAGAAAGCAACCAATGTTTACATATATAGTGTAAAGTAA 2333
QY 301 TCTGGAATATCTTCTGCTGTGATCCTCCAGTAAATATCCGGATTTACCTTAATGTAAA 360
DB 2334 TCTGGAATATCTTCTGCTGTGATCCTCCAGTAAATATCCGGATTTACCTTAATGTAAA 2393
```

```
QY 361 ACTGTTTCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATTTGCCCTCCATTTGTGG 420
DB 2394 ACTATCCATCATATTTCAAGGTCAAAACCCCTCATGCACAGGGGATTCGCCCTTCATTTATGG 2453
QY 421 GGGGCAATTTTCTGTATGATCGGGTTCCTCTCAACAATGTACCGAGGCAAGGCTCTTC 480
DB 2454 GGAGCAATTTTCTGTATGATCGCATTTGCCCTCCACAACAATGTACCGAGGCAAGGCTCTTC 2513
QY 481 ACTCAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGTTCAAGAAATGATTTTCT 540
DB 2514 ACTCAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGTTCAAGAAATGATTTTCT 2573
QY 541 AGGCAAGGACAAAGGTTATCGTCAATGAACTTGACCTCCACCAATATAATATTGGACAAGC 600
DB 2574 CGGCAAGGACAAAGGTTATCGTCAATGAACTTGACCTCCACCAATATAATATTGGACAAGT 2633
QY 601 AGCAATGAAGCGCAGAGAAATGATACGGGATGTTTTTGGCATCTCCCAAGAAATACAACTCC 660
DB 2634 AGTAAACGGAACGCAACGAAATGACACTGGATGTTTTCGGCGCTCTTCAAGAAATACAACTCT 2693
QY 661 ACAAACAATCAAAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACTCCG 720
DB 2694 ACAAAGAACCAAAACATGTGCTCCGTCCAAATATACCTCCACCATCTGCCACAGCCCGTCCG 2753
QY 721 AGCAATTCACCTACAAAATATCTCAAAATTAATACTGCTAAATCTGGAACCTATGAAACCAAGT 780
DB 2754 GAGATCAAACTCAACAGCACCCCAACTGATGCGCACCAAACTCAATACCAAGGACCAAGC 2813
QY 781 AGCAGCATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAACAGGGGGCCCAACACA 840
DB 2814 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCGAGAGAACGAGAAACCCCAACA 2873
QY 841 ACTCTTAATGTAGTCACTGACACAGAAACAATCGTCAACAATATTGTCCACATCTCTCACTA 900
DB 2874 ACTTCTGATGGGTCACCAAGCAAGGGCTTTTCATCAACATGCGCACCCCTCCCTCCTCA 2933
QY 901 CATCCAAGCACCTCAACAACATGAGCAAAAACAGTAGTACGAATCTTCCCGACATGCTGTAACT 960
DB 2934 CAACCAAGCACGCGCACAGGAGGAGGAAACAACAACCAATTTCCCAAGATGCTGTGACT 2993
QY 961 GAGCACAATGGAAACCGACCCCAACAACAACAAGAGCGTCTCTCAACAATATCTAATACA 1020
DB 2994 GAACTAGACAAAAATAACACAACCTGCAACACCGTCCATGCGCCCTCTATAAACACTACCA 3053
QY 1021 ACTCCCACTTATACATCTCTCAAGTACAACTCTAGTACTCTTCCCTCCCAACCGCAAC 1080
DB 3054 ATCTCTACTAACAACACCTCCAAACACAACTTTCAGCACTCTCTCTGCAACCAATTAACAAC 3113
QY 1081 ATCAACCAATATGATACACAACGTCGAACCTAGCAGAAAACGCAAAACCAATGCTCAGTTG 1140
DB 3114 ACCCAATGACACACACAGAGCACAACTCACTGAAATGAGCAAAACCAAGTCCCTCCCTCG 3173
QY 1141 AACACAACCTTAGATCCAAAGAAAAATCCCAACAGAGCAAGACACCAACAGCAACACC 1200
DB 3174 ATAAACAACCTTCCTCCAAACGGAATATCCCAACAGCAAAAGAGCAGCAGCAGCAAAAAA 3233
QY 1201 AACATCATGATGAGCAGATCAGATATACAGAGCAAAACCCCAACAAATTTCTTCCGGAT 1260
DB 3234 GGCCCCGCAACAGGGCACCAACACGACAAATGAGCAATTTCCAGCTCTCCCTCCCTCCAC 3293
QY 1261 TCTAGTCCGCAACACCGCCCTCTTATATATCTTTTAAAGAAAGAACCGATCGATCCTCTGGAGG 1320
DB 3294 CCCAGCTCGATGCAACAACATCTTGTATATTTTCAGAAAGAGCAAGTATCTCTGGAGG 3353
QY 1321 GAAGCGCAGATGTTCCCTTTTCTGATGGGTTAATAAATGCTCCATTTGATTTTGACCCA 1380
DB 3354 GAAGCGCAGATGTTCCCTTTTCTGATGGGTTAATAAATGCTCCATTTGATTTTGACCCA 3413
QY 1381 GTTCCAAATACAAAACAACATCTTTGATGAATCTCTAGTCTGTGGTCCCTCGCTCAGGAA 1440
DB 3414 GTTCCAAATACAAAACAACATCTTTGATGAATCTCTAGTCTGTGGTCCCTCGCTCAGGAA 3473
```

Qy 1441 GATCAATGCTCCCGCAATATAGTTAACTTTATCTTATTTCTTAATATAATGAG 1500
Db |||||
Qy 3474 GATCAATGCTCCCGCAATATAGTTAACTTTATCTTATTTCTTAATATAATGAG 3533
Db |||||
Qy 1501 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGATGAGAGTTAAGAAATTTGGAGC 1560
Db |||||
Qy 3534 AACACTGCTACTCTGGAGAAATGAGAAATGATTTGATGAGAGTTAAGAAATTTGGAGC 3593
Db |||||
Qy 1561 GTTCAGAGAGATGACCTGGCGAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 1620
Db |||||
Qy 3594 GTTCAGAGAGATGACCTGGCGAGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 3653
Db |||||
Qy 1621 GAAGGACTTTACACTGCTGTTTTTAATTAATAATCAAAACAAATTTGGTCTGCGAGTTGAGG 1680
Db |||||
Qy 3654 GAAGGACTTTACACTGCTGTTTTTAATTAATAATCAAAACAAATTTGGTCTGCGAGTTGAGG 3713
Db |||||
Qy 1681 CGTCTAGCCAATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACCTGAGGAA 1740
Db |||||
Qy 3714 CGTCTAGCCAATCAAACTGCCAAATCCTTGGAACTCTTATTGAGAGTCACAACCTGAGGAA 3773
Db |||||
Qy 1741 AGAATCATCTCTTAATCAATAGACATGCTATTGATGCTTCTACTCAAAAGATGGGAGGA 1800
Db |||||
Qy 3774 AGAATCATCTCTTAATCAATAGACATGCTATTGATGCTTCTACTCAAAAGATGGGAGGA 3833
Db |||||
Qy 1801 ACATGCAAGGCTTGGACCTGATTTGATCGGATAGAGACTTGTCCAAAATATT 1860
Db |||||
Qy 3834 ACATGCAAGGCTTGGACCTGATTTGATCGGATAGAGACTTGTCCAAAATATT 3893
Db |||||
Qy 1861 TCAGAGCAAAATGACCAAAATTAATAAGGAGCAACAAAGAGGGGACTGGTTGGGGCTG 1920
Db |||||
Qy 3894 TCAGAGCAAAATGACCAAAATTAATAAGGAGCAACAAAGAGGGGACTGGTTGGGGCTG 3953
Db |||||
Qy 1921 GGTGGTAAATGGTGACATCCGACTCGGGTGTTCTTACTAACTTGGGCAATTTGCTACTA 1980
Db |||||
Qy 3954 GGTGGTAAATGGTGACATCCGACTCGGGTGTTCTTACTAACTTGGGCAATTTGCTACTA 4013
Db |||||
Qy 1981 TTATCCATAGCTGCTTGAATGCTATCTCTGATTTGCTGATCTTACTAAATATATC 2040
Db |||||
Qy 4014 TTATCCATAGCTGCTTGAATGCTATCTCTGATTTGCTGATCTTACTAAATATATC 4073
Db |||||
Qy 2041 GGATAA 2046
Db |||||
Qy 4074 GGATRA 4079
Db |||||

RESULT 4

US-10-267-322-1
; Sequence 1, Application US/10267322
; Publication No. US20030152590A1
; GENERAL INFORMATION:
; APPLICANT: United States Army Medical Research Institute of
; APPLICANT: Infectious Diseases
; APPLICANT: Hevey, Michael C.
; APPLICANT: Negley, Diane L.
; APPLICANT: Pushko, Peter
; APPLICANT: Smith, Jonathan F.
; APPLICANT: Schmaljohn, Alan L.
; TITLE OF INVENTION: Marburg Virus Vaccines
; FILE REFERENCE: Army 143D
; CURRENT APPLICATION NUMBER: US/10/267,322
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: US 09/336,910, and US 60/091,403
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Word, Microsoft Office 97, IBM compatible
; SEQ ID NO 1
; LENGTH: 11460
; TYPE: DNA
; ORGANISM: Marburg Virus
US-10-267-322-1

Query Match 74.0%; Score 1514.8; DB 16; Length 11460;
Best Local Similarity 83.8%; Pred. No. 0;

Matches 1714; Conservative 0; Mismatches 332; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTTCTGATTTAGTCTCATTTTATCAAAAGTATAAAACTCTCCCT 60
Db |||||
Qy 5940 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTAAATCAAGGACAAATAATCTCCCC 5999
Db |||||
Qy 61 GTTTTAGAATAATGCTAGTAACAGCAACCTCAAGATGTAGATTTCAGTGTGCTCCGGAAAC 120
Db |||||
Qy 6000 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGATTCGGTATGCTCCGGAACT 6059
Db |||||
Qy 121 CTCCAAAAGACAGAAAGATGTTCTATCTGATGGGATTTACACTGAGTGGCAAAAAGTTGCT 180
Db |||||
Qy 6060 CTCGAGAGACAGAGACGTCATCTGATGGGATTTACACTGAGTGGCAAAAAGTTGCT 6119
Db |||||
Qy 181 GATTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAAGACAGGTGTTCTCCCAAGAAC 240
Db |||||
Qy 6120 GATTCCCTTTTGAAGCATCTAAACGATGGGCTTTTCAAGACAGGTGTTCTCCCAAGAAC 6179
Db |||||
Qy 241 GTTGGATATACGAGAGAGAGAGCCAAACATGTTTACAATATAAGTGTACAGACCTT 300
Db |||||
Qy 6180 GTTGGATATACGAGAGAGAGAGCCAAACATGCTTACAATATAAGTGTACAGACCTT 6239
Db |||||
Qy 301 TCTGGAATAATCTTGTCTGCTGATCTCTCCAGTAATATCCGCGATTACCTTAAATGTAAA 360
Db |||||
Qy 6240 TCTGGAATAATCTTGTCTGCTGATCTCTCCAGTAATATCCGCGATTACCTTAAATGTAAA 6299
Db |||||
Qy 361 ACTGTTTCATCATATTTCAAGGTCAAAACCTCATGCAACAGGGGATTTGCCCTCCATTGTTGG 420
Db |||||
Qy 6300 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCAACAGGGGATTTGCCCTCCATTGTTGG 6359
Db |||||
Qy 421 GGGGCATTTTCTGATGATCGGGTTCCTTACAAATATACCGAGGCAAGGTCTTC 480
Db |||||
Qy 6360 GGAGCATTTTCTGATGATCGCATTTGCCCTCCACAAATATGTAACCGAGGCAAAAGTCTTC 6419
Db |||||
Qy 481 ACTGAAGGAATATAGCAGCTATGATTTGTTAATGAACAGTTCACAGATGATTTTTTCT 540
Db |||||
Qy 6420 ACTGAAGGAATATAGCAGCTATGATTTGTTAATGAACAGTTCACAGATGATTTTTTCT 6479
Db |||||
Qy 541 AGSCAAGGACAAAGGTATCTGTCATGAACTTGACCTCCCAATAAATATTTGGACAAAGC 600
Db |||||
Qy 6480 CGGCAAGGACAAAGGTATCTGTCATGAACTTGACCTCCCAATAAATATTTGGACAAAGT 6539
Db |||||
Qy 601 AGCAATGAAACGCAAGAGAAATGATA CGGGATGTTTGGGATCTCTTCCAAAGAAATACAACTCC 660
Db |||||
Qy 6540 AGTAAACGGAACGCAAAACGAATGACATCGGTGATGTTTCCGGCGCTCTTCAAGAAATACAACTCT 6599
Db |||||
Qy 661 ACAAAACATCAACATGCTCCCTCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACCTCCG 720
Db |||||
Qy 6600 ACAAAAGAACAAACATGCTGCTCCGTCCTCAAAATACCTCCACACCTGCCCCACAGCCCGTCCG 6659
Db |||||
Qy 721 AGCATTCACCTCTACAAATATCTCAAAATTAATCTGCTTAAATCTGGAACCTATGAAACCCAAAGT 780
Db |||||
Qy 6660 GAGATCAAACTCACAGCACCCCACTGATGCCCAAACTCAATACCAAGGACCCCAAGC 6719
Db |||||
Qy 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGAGAGAACAGGGGCCCCACACA 840
Db |||||
Qy 6720 AGTGATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGAGAGAACGAGAACCCCAACA 6779
Db |||||
Qy 841 ACTCTTAATGTAGTCACTGAACAGAAACATCTGTCACAAATATTTGTCACCTCTTCACTA 900
Db |||||
Qy 6780 ACTTCTGATCGGTCACCAAGCAAGGGCTTTTCAATCAAAATGCCCACCTCCCTCAACA 6839
Db |||||
Qy 901 CATCCCAAGCACCTCACAAATGAGCAAAACAGTACGAATCTCTCCGACATGCTGTAACT 960
Db |||||
Qy 6840 CAACCAAGCAGCCACAGAGAGGAGGAGAACACACAAACCAATTTCCCAAGATGCTGTGACT 6899
Db |||||
Qy 961 GAGCACAATGGAACCGGACCCCAACAAACAGCAACAGCTCTCTTCAACAATATTAATACA 1020
Db |||||
Qy 6900 GAACTAGACAAAATAACACAACTGCAACACCGTCCATGCCCTCATATAACACTTACCACA 6959
Db |||||
Qy 1021 ACTCCCAAGCACCTCACAAATGAGCAAAACAGTACGAATCTCTCCGACATGCTGTAACT 1080
Db |||||
Qy 6960 ATCTCTACTAACCAACCTCCCAACACAACTTCAAGCACTCTCTCTGCAACCAATTAACAAAC 7019
Db |||||

```
QY 1081 ATCAACAATAATGATACACAACTGTAAGTAACTAGCAGAAACGAAACCAACCAATGCTCAGTTG 1140
Db 7020 ACCACCAATGACAAACACACAGAGCAACATCACTGAAATGAGCAACCAAGTGCCTCCCTCG 7079
QY 1141 AACCAACTCTAGATCCCAACAGAAATCCCAACAGAGCAAGACCAACCAAGCACAACC 1200
Db 7080 ATAACAACCTGCTCCCAACGGAATCCCAACAGCAAGAGCACCAGCAGCAAAAAA 7139
QY 1201 AACATCATCATGACGACATCAGATATAACAAGCAAAACACCCACAAATTTCTTCCGGAT 1260
Db 7140 GGCCCCGCCACACAGGACCAACACAGCAAAATGAGCATTTTCAACAGTCTCCCCCACC 7199
QY 1261 TCTAGTCCGCAACACCGCCCTCTATATACCTTTAGAAAGAAACGATCGATCCTCTGGAGG 1320
Db 7200 CCCAGCTCGATGCAACAACATCTTGTATATTTCAAGAAAGCGAAGTATCTCTGGAGG 7259
QY 1321 GAAGGCGCATGTTCCCTTTCTTGATGGTTTAATAATGCTCCAAATTTGATTTTGACCCA 1380
Db 7260 GAAGGCGCATGTTCCCTTTCTTGATGGTTTAATAATGCTCCAAATTTGATTTTGACCCA 7319
QY 1381 GTTCCAAATACAAAACCAATCTTTGATGAATCCTCTAGTTCTGTGTCCTCGCTGAGGAA 1440
Db 7320 GTTCCAAATACAAAACCAATCTTTGATGAATCCTCTAGTTCTGTGTCCTCGCTGAGGAA 7379
QY 1441 GATCAACATGCTCCCTCCCAATATTAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG 1500
Db 7380 GATCAACATGCTCCCTCCCAATATTAGTTTAACTTTATCTTATTTTCTTAATATAAATGAG 7439
QY 1501 AACCTGCTACTCTGGAGAAATGAGAAATGAGAAATGATGATGATGATGATGATGATGATGATG 1560
Db 7440 AACCTGCTACTCTGGAGAAATGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 7499
QY 1561 GTTCAGGAGGATGACTCGCGCAGGCTCAGTTGCGATACCGTTTTTTTGGCCCTGGAAAT 1620
Db 7500 GTTCAGGAGGATGACTCGCGCAGGCTCAGTTGCGATACCGTTTTTTTGGCCCTGGAAAT 7559
QY 1621 GAAGGACTTTTACATGCTGTTTTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1680
Db 7560 GAAGGACTTTTACATGCTGTTTTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 7619
QY 1681 CGTCTAGCAATCAAACTGCGAAATCTCTTGGAACTCTTATTTGAGAGTCAACAATGAGGAA 1740
Db 7620 CGTCTAGCAATCAAACTGCGAAATCTCTTGGAACTCTTATTTGAGAGTCAACAATGAGGAA 7679
QY 1741 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGGGGA 1800
Db 7680 AGAATCTCTCTTAATCAATAGACATGCTATTGACTTTCTACTCACAAGATGGGGGGA 7739
QY 1801 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAAGACTTGTCCAAAATAATT 1860
Db 7740 ACATGCAAAAGTCTTGGACCTGATTTGTCATCGGGATAGAAGACTTGTCCAAAATAATT 7799
QY 1861 TCAGAGCAAAATGACAAATTAATAAGGACGAAACAAAAGAGGGGACTGGTTGGGGCTG 1920
Db 7800 TCAGAGCAAAATGACAAATTAATAAGGACGAAACAAAAGAGGGGACTGGTTGGGGCTG 7859
QY 1921 GGTGCTAAATGGTGACATCCGACTGGGGTGTCTTACTTAATTTGGGCAATTTGCTACTA 1980
Db 7860 GGTGCTAAATGGTGACATCCGACTGGGGTGTCTTACTTAATTTGGGCAATTTGCTACTA 7919
QY 1981 TTATCCATAGTCTCTTGATGCTCTATCTCTGATTTTGTCTGATCTTTTACTAAATATATC 2040
Db 7920 TTATCCATAGTCTCTTGATGCTCTATCTCTGATTTTGTCTGATCTTTTACTAAATATATC 7979
QY 2041 GGATAA 2046
Db 7980 GGATAA 7985
```

RESULT 5
US-10-066-506A-9
; Sequence 9, Application US/10066506A

```
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 9
; LENGTH: 2051
; TYPE: DNA
; ORGANISM: Marburg virus strain Musoke
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus Glycoprotein 1 and Marburg
; OTHER INFORMATION: Glycoprotein 2
US-10-066-506A-9
```

Query Match 72.6%; Score 1485.6; DB 15; Length 2051;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1705; Conservative 0; Mismatches 329; Indels 12; Gaps 1;

```
QY 1 ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAAACTCTCCCT 60
Db 10 ATGAAGACCATATGTTTCTTATCAGTCTTATCTTAATCAAGGACAAAANAATCTCCCC 69
QY 61 GTTTTAGAATTTGCTAGTAAACAGCCAACTCAAGATGTAGATTGATGTGTCCTCGGAACC 120
Db 70 ATTTTAGAGATAGCTAGTAAATCAACCCCAAAATGTGGATTGGTATGTCCTCGGAAC 129
QY 121 CTCCAAAGACAGAGATGTTTCATCTGATGGGATTTACATGATGGCGATACCTTAAAGATTGCT 180
Db 130 CTCCAAAGACAGAGATGTTTCATCTGATGGGATTTACATGATGGCGATACCTTAAAGATTGCT 189
QY 181 GATTCCCTTTTGAAGCATCTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 190 GATTCCCTTTTGAAGCATCTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 249
QY 241 GTTCAGTATACGGAAGGAGAAAGCCAAACATGTTTCAATATAAGTGTAAACGACCT 300
Db 250 GTTCAGTATACGAGGGGAGAGCCAAACATGCTACAATAAAGTGTAAACGATCCC 309
QY 301 TCTGGAATTCCTTGTCTGGATCCCTCCAGTAAATATCCCGATTAACCTTAATGTAAA 360
Db 310 TCTGGAATTCCTTGTCTGGATCCCTCCAGTAAATATCCCGATTAACCTTAATGTAAA 369
QY 361 ACTGTTCAATATTCAGGTCAAAACCTCATGCACAGGGGATGCGCTCCATTTGTGG 420
Db 370 ACTATCAATATTCAGGTCAAAACCTCATGCACAGGGGATGCGCTCCATTTATGG 429
QY 421 GGGCATTTTCTTGTATGATCGGTTGCTCTCAACAATGTACGAGCAAGGCTTTC 480
Db 430 GGAGCATTTTCTTGTATGTCGCAAT-----ATGACCGAGGCAAGTCTTC 477
QY 481 ACTGAAGAAATATAGCAGCTATGATTTGTTAATGAAGACAGTTCAAGATGATTTTTTCT 540
Db 478 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATGAAGACAGTTCAAGATGATTTTTTCTG 537
QY 541 AGGCAAGGACAAAGTTATCGTCACATGAACTTGACCTCCACAAATAAATATTTGGACAAGC 600
Db 538 CGGCAAGGACAAAGTTATCGTCACATGAACTTGACCTCCACAAATAAATATTTGGACAAGT 597
QY 601 AGCAATGAACGACAGAAATGATACGGGATGTTTTTGGCATCTCTCCAAAGATACAACTCC 660
Db 598 AGTAACGGACGCAACGAAATGACATGGAATGTTTGGCGCTCTTCAAGAAATACAAATCT 657
QY 661 ACAACCAATCAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG 720
Db 658 ACAAGAACCAACATGCTGCTGTCCTCAAAATATCCTCCACCTGCCCCACAGCCCGTCCG 717
```

Qy	721	AGCAATCACTCTACAAATAC	CTCAAAATTAATACTGCTAAATCTGGA	ACTATGAACCCCAAGT	780
Db	718	GAGATCAAACTCACAAGCAC	CCCAACTGATGCGCACCAAACTCAAT	ACCAAGGACCCCAAGC	777
Qy	781	AGCCACGATGAGGACCTTAT	GATTTTCGGGCTCAGGATCTGGAA	CAGGGGCCCAACACA	840
Db	778	AGTGATGATGAGGACCTCG	CAACATCCGGCTCAGGGTCCGGAG	ACGAGAACCCCAACA	837
Qy	841	ACTCTTAATGTAGTCACTG	AACAGAAACAATCTGTCAACAATAT	TGTCCACTCTCTCACTA	900
Db	838	ACTTCTGATGGGTCAACA	AGCAAGGGCTTTCATCAACAAT	TGCCACCACTCTCTCTCA	897
Qy	901	CATCCAAGCACCTCACAAC	ATGAGCAAAAACAGTACGAATCT	CTTCCCGACATGCTGTACT	960
Db	898	CAACCAAGCACGCCACAG	CAGAGGAGGAAAACAACAACCA	ATTCCTCCAAAGATGCTGTGACT	957
Qy	961	GAGCACAAATGGAA	CCGACCCAAACAACAACCAAGCA	ACGCTCTCTCAACAATACTAT	1020
Db	958	GAATAGACAAAATAACA	AACTTGCACACCGTTCATGCCCC	CTCATTAACACTACCA	1017
Qy	1021	ACTCCCACCTATAACACT	CTCTCAAGTACAACCTCAGTACT	CTCTTCCCTCCAAACCCGCAAC	1080
Db	1018	ATCTCTACTRAACAAC	CTCCAAACAACATTCAGCACTCT	CTCTGTGCACCATTACAAAC	1077
Qy	1081	ATCACCAATAATGATACA	CAACAGTGAACTAGCAGAAAGCG	AAACCAATGCTCAGTTG	1140
Db	1078	ACCAACCAATGACAACA	CAGAGCACAATCACTGAAAAT	GAGCAAAACGAGTGCCTCTCG	1137
Qy	1141	AACAACACTCTAGATCCA	ACAGAAAAATCCCAACACAGACA	AGACACCAACAGACAAC	1200
Db	1138	ATAACAACCTTGCTCC	NAACGGGAATCCCAACACAGAA	AGACACAGAGCAAGAAAA	1197
Qy	1201	AACATCATATGACGACAT	CAGATATAACAAAGCAAAAC	CCCCCAAAATTTCTTCGGAT	1260
Db	1198	GGCCCCGCCACAAGGG	CACCAACAGCAAAATGAGCA	ATTTTCCAGCTCTCTCCCCCACC	1257
Qy	1261	TCTAGTCGCAACCCG	CCCTCTATATCTTTTAAAGAA	AGAACGATCGATCTCTTGAGG	1320
Db	1258	CCCAGCTCGATCGACA	CAACTCTTGATATTTTCAGAA	AGAACGATCGATCTCTTGAGG	1317
Qy	1321	GAAGCGACATGTTCC	CTTTTCTGGATGGGTAAATAA	TGCTCCAAATGCTTTCGACCCA	1380
Db	1318	GAAGCGACATGTTCC	TTTTCTGGATGGGTAAATAA	TGCTCCAAATGCTTTCGACCCA	1377
Qy	1381	GTTCCAAATACAAAAA	CAACTCTTTGATGAATCTCTAG	TGTTGTCGCTCGGCTGAGAA	1440
Db	1378	GTTCCAAATACAAAAA	CAACTCTTTGATGAATCTCTAG	TGTTGTCGCTCGGCTGAGAA	1437
Qy	1441	GATCAACATGCTCCCC	CAATATAGTTTAACTTTTATCT	TATTTTCTTAATAATAATGAG	1500
Db	1438	GATCAACATGCTCCCC	CAATATAGTTTAACTTTTATCT	TATTTTCTTAATAATAATGAG	1497
Qy	1501	AACACTGCTACTCTG	GAGAAAAATGAGAAATGATGTG	ATGCAGAGTTTAAAGAAATTTGGAGC	1560
Db	1498	AACACTGCTACTCTG	GAGAAAAATGAGAAATGATGTG	ATGCAGAGTTTAAAGAAATTTGGAGC	1557
Qy	1561	GTTCAAGGAGATGAC	CTGGCCGACAGGCTCAGTTG	GAATACCGTTTTTGGCCCTGGAATT	1620
Db	1558	GTTCAAGGAGATGAC	CTGGCCGACAGGCTCAGTTG	GAATACCGTTTTTGGCCCTGGAATT	1617
Qy	1621	GAAGGACATTACATG	CTGTTTAAATTAATAAATCAAA	ACAATTTGGTCTGCAGGTTGAGG	1680
Db	1618	GAAGGACATTACATG	CTGTTTAAATTAATAAATCAAA	ACAATTTGGTCTGCAGGTTGAGG	1677
Qy	1681	CGTCTAGCCAATCAA	CTGCCCCAATCTTGTGAACTCT	TATTTGAGAGTCAACAATGAGAA	1740
Db	1678	CGTCTAGCCAATCAA	CTGCCCCAATCTTGTGAACTCT	TATTTGAGAGTCAACAATGAGAA	1737
Qy	1741	AGAACATTTCTCCT	TAATCAATAGACATGCTATTG	ACTTCTACTCAACAGATGGGGAGGA	1800
Db	1738	AGAACATTTCTCCT	TAATCAATAGACATGCTATTG	ACTTCTACTCAACAGATGGGGAGGA	1797

QY	1801	ACATGCAAAAGTCTTGGACCTGATTTGTTCATCGGGATAGAAGACTTGTCCAAAAATATT	1861
DB	1798	ACATGCAAAAGTCTTGGACCTGATTTGTTCATCGGGATAGAAGACTTGTCCAAAAATATT	1857
QY	1861	TCAGAGCAAAATTCGACCAAAATTTAAAGAGGACGAAACAAAGAGGAGGACTGTTGGGGTCTG	1920
DB	1958	TCAGAGCAAAATTCGACCAAAATTTAAAGAGGACGAAACAAAGAGGAGGACTGTTGGGGTCTG	1917
QY	1921	GGTGTTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCATTTTGTCTACTA	1980
DB	1918	GGTGTTAAATGGTGGACATCCGACTGGGGTGTCTTACTAACTTGGGCATTTTGTCTACTA	1977
QY	1981	TTATCCATAGCTGTCTTGATTTGCTCTATCTCTGATATTTGTCTGATCTTTACTAAATATATC	2040
DB	1978	TTATCCATAGCTGTCTTGATTTGCTCTATCTCTGATATTTGTCTGATCTTTACTAAATATATC	2037
QY	2041	GGATAA 2046	
DB	2038	GGATAA 2043	
RESULT 6			
US-10-811-353-2			
; Sequence 2, Application US/10811353			
; Publication No. US20050112098A1			
; GENERAL INFORMATION:			
; APPLICANT: McCray, Paul B.			
; APPLICANT: Sanders, David A.			
; APPLICANT: Jeffers, Scott A.			
; APPLICANT: Davidson, Beverly L.			
; APPLICANT: Sinn, Patrick L.			
; TITLE OF INVENTION: PSEUDOTYPED VIRUSES AND METHODS FOR THEIR USE			
; FILE REFERENCE: 290.00670120			
; CURRENT APPLICATION NUMBER: US/10/811,353			
; CURRENT FILING DATE: 2004-03-26			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: PatentIn version 3.2			
; SEQ ID NO 2			
; LENGTH: 2046			
; TYPE: DNA			
; ORGANISM: Marburg virus glycoprotein			
US-10-811-353-2			
Query Match 72.0%; Score 1473.2; DB 21; Length 2046;			
Best Local Similarity 82.5%; Pred. No. 0;			
Matches 1688; Conservative 0; Mismatches 358; Indels 0; Gaps 0;			
QY	1	ATGAAGACCATATATTTTCTGATTAGTCTCATTTTAAATCCAAAGTATAAAAACTTCCTCT	60
DB	1	ATGAAGACCATATGCTCTTTATCAGTCTTATCTTAATCCAAAGGATAAAAACTTCCTCT	60
QY	61	GTTTTAGAAATTCGTAGTAAACAGCCCAACCTCAAGATGTAGATTTCAGTGTGCTCCGGAACT	120
DB	61	ATTTTAGAGATAGCTAGTAAACATCAACCCCAAATGTGGATTCGGTATGCTCCGGAACT	120
QY	121	CTCCAAAGACAGAAGATGTTTCATCTCATGCGGATTTACCTGAGTGGGCAAAAAGTTGCT	180
DB	121	CTCCAGAGACAGAAGATGTTTCATCTCATGCGGATTTACCTGAGTGGGCAAAAAGTTGCT	180
QY	181	GATTCCTCTTTGGAAGCATCTAAACGATGGGCTTTTCAGGACAGTGTCTCTCCCAAGAAC	240
DB	181	GATTCCTCTTTGGAAGCATCTCAACGAGTGGGCTTTTCAGGACAGTGTACCTCCCAAGAAC	240
QY	241	GTTTCAGTATACGGAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGACCT	300
DB	241	GTTTCAGTATACGGAGGAGAGAGCCAAACATGTTTACAATATAAGTGTAAACAGATCCC	300
QY	301	TCTGAAAATCCTTGCTGCTGGATCTCTCCCAAGTAAATATCCCGCATTTACCTAAATGTAAA	360
DB	301	TCTGAAAATCCTTGCTGCTGGATCTCTCTACCAACATCCGTGACTATCTTAATGTCAA	360
QY	361	ACTGTTTCATCATATTCAGAGTCAAAACCTCATGACAGGGGATTCGCCTCATTTGTGG	420


```
Db 361 ACTATCCATCATATTTCAAGGTCATAAAACCCCTCATGTCGCAAGGGATGCGCCCTCCATTTGTGG 420
Qy 421 GGGGCAITTTTCTTGATGTCGCGTTCCTCTCAACAACATGTATCCGAGGCAAGTCTTTC 480
Db 421 GGAGCAITTTTCTTGATGTCGCAITTCCTCCACAACAATGTATCCGAGGCAAGTCTTTC 480
Qy 481 ACTGAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGATTCACAGAATGATTTTCT 540
Db 481 ACTGAAGGGAACATAGCAGCTATGATTTGTTAATAAGACAGATTCACAGAATGATTTTCTCG 540
Qy 541 AGGCAAGGACAGGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTTGACAGC 600
Db 541 AGGCAAGGACAGGTTATCGTCACATGAACCTTGACCTCCACCAATAAATATTTGACAGC 600
Qy 601 AGCAATGAACGCGAGAGAAATGATACGGGATGTTTGGCATCCTCCAGAATATACAATCC 660
Db 601 AACCAATGAACACAAACGAATGACACTGGATGCTTCGGTGCTCTCAAGATATCACTCC 660
Qy 661 ACAAACAATCAAAACATGCCCTCCATCTCTTAAACCTCATCCCTGCCGCCACAGTACTCCG 720
Db 661 ACGAAGAAATCAAAACATGCTCCGTCCAAATACCCCTCACCACTGCCGCCACAGCCGCTCA 720
Qy 721 AGCATTCATCTACAAATACTCAAAATTAATCTGCTAAATCTGGAACCTATGAACCAAGT 780
Db 721 GAGATCAAAACCCACAAGCACCCCAACTGATGCCACCACTCAACACCAAGACCCCAAC 780
Qy 781 AGCGACGATGAGGACCTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db 781 AATGATGATGAGGACCTCATAACTCCGTTTCAGGGTCGGAGAACAGGACCCCTATACA 840
Qy 841 ACTCTTAATGATGACCTGAAACAGAAAACATCGTCAACAATATGTCACCTCTCTTCACTA 900
Db 841 ACTTCAGATGGGTCACCTAAGCAAGGGCTTTTCATCAACAATGCCACCACTCCCTCAACA 900
Qy 901 CATCAAGCACTCAACATGAGCABAACAGTACGATCTCTCCGACATGCTGTACT 960
Db 901 CAACCAAGCAGCCACAGCAAGAGGAAACACACAGACCACTATCCCAAGGTAAGTGTACT 960
Qy 961 GAGCAAAATGGAACCGACCCCAACACACACAGCAAGCTCCTCAACAATACTAATAACA 1020
Db 961 GAACCCACAAACCAACACACAGCGCACACACCGTCCATGCGCCGCCACCAACCACTGCA 1020
Qy 1021 ACTCCCACTATACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCTCCAAACCCGCAAC 1080
Db 1021 ATCTCTACTAAACACACCTCCAAAGAACAACTTCAGCACCTCTCTGTATCACTACAAAC 1080
Qy 1081 ATCACCATATGATACACACGTAAGTACTAGCAGAAAGGACAAACCAATGCTCAGTTG 1140
Db 1081 ACCACCAATAGCACACAGACGACAGCCACTGAAAATGAAACAAACAGTGCCCCCTCG 1140
Qy 1141 AACCAACTCTAGATCCCAACAGAAAATCCCAACAGGACAGACACCAACAGCACACACC 1200
Db 1141 AAACAAACCTCGCTCCAAAGGAATCTTACCAGCAGAAAGACACTTAAACACAGAA 1200
Qy 1201 AACATCATATGACGACATCATAGATATAACAGCAAAACACCCACAAATCTTCTCCGAT 1260
Db 1201 GSCCCCAACCAACCGCACCAAAATATGACAAATGGCATTTAAACGATCCCTCCGCCACC 1260
Qy 1261 TCTAGTCCGACAAACCGCCCTCCTATATACCTTATAGAAAGACGATCGATCTCTGGAG 1320
Db 1261 CCCAACCCGACACAAACATCTTGTATATTTTCAAGAAAGAACGAAATCTCTCTGGAG 1320
Qy 1321 GAAGCGCATGTTCCCTTTCTGATGGGTTAATAAATGCTCCAAATGATTTTGGACCA 1380
Db 1321 GAAGCGCATGTTTCTTTCTGACGGGTTAATAAATGCTCCAAATGATTTTGGATCA 1380
Qy 1381 GTTCCAAATACAAAACAAATCTTTGATGAATCTCTAGTCTGTGGGCTCCGGCTGAGAA 1440
Db 1381 GTTCCAAATACAAAGACGATCTTTGATGAATCTTCTAGTCTGTGGGCTTCGGCTGAGAA 1440
Qy 1441 GATCAACATGCTCCGCCAAATATAGTTTAACTTTATCTTATTTCTTAAATATAAATGAG 1500
Db 1441 GATCAACATGCTCCGCCAAATATCAGTTTAACTTTATCTTATTTCTTAAATATAAATGAA 1500
```

```
Qy 1501 AACACTGCCTACTCTGAGAGAAATGAGAAATGATTTGTATGTCAGAGTTAAAGAAATTTGGAGC 1560
Db 1501 AACACTGCCTACTCTGGAGAGAAATGAGAAACGATTTGTATGTCAGAGTTAAAGAAATTTGGAGC 1560
Qy 1561 GTTCAGAGGATGACCTGGCCGACAGGCTCAGTTTGGATACCGTTTGTGGCCCTGGAAT 1620
Db 1561 GTTCAGAGGATGACCTGGCCGACAGGCTCAGTTTGGATACCGTTTGTGGCCCTGGAAT 1620
Qy 1621 GAAGCACTTTACACTGCTGTTTAAATTAATAAATCAAAACAAATTTGGTCTGCAAGTTGAGG 1680
Db 1621 GAAGCACTTTATCTGCTGTTTAAATTAATAAATCAAAACAAATTTGGTCTGCAAGTTGAGG 1680
Qy 1681 CGTCTAGCCAATCAAACTGCCAAATCCTTGGAACTCTTATTTAGAGTCACAACCGAGAA 1740
Db 1681 CGTCTAGCCAATCAAACTGCCAAATCCTTGGAACTCTTATTTAGAGTCACAACCGAGAA 1740
Qy 1741 AGACATTCCTTTAATCAATAGACATGCTATTGACATTTCTACTCACAGATGGGAGGA 1800
Db 1741 AGGACATTTCTTTAATTAATAGACATGCCATTTGACTTCTACTCACAGATGGGAGGA 1800
Qy 1801 ACATGCAAAAGTCTTGGACCTGATTTGTTCATCGGATAGAGACTTGTCCAAAATAT 1860
Db 1801 ACATGCAAAAGTCTTGGACCTGATTTGTTCATCGGATAGAGACTTGTCCAGGAATAT 1860
Qy 1861 TCAGAGCAAAATTTGACCAAAATTTAAAAAGGACGAAACAAAAAGGGGACTGTTGGGCTCTG 1920
Db 1861 TCGGAAACAAATTTGACCAAAATTTAAAAAGGATGAAACAAAAAGGGGACTGTTGGGCTCTA 1920
Qy 1921 GGTGTTAAATGGTGACATCCGACTGGGTGTTCTTACTACTTGGGCAATTTGCTACTA 1980
Db 1921 GGTGTTAAATGGTGACATCCGACTGGGTGTTCTTACTACTTGGGCAATTTGCTACTA 1980
Qy 1981 TTATCCATAGCTGTCTTGATTTGCTATCTATCTGATTTGTCTGATCTTTACTAAATATATC 2040
Db 1981 TTATCCATAGCTGTCTTGATTTGCTATCTATCTGATTTGTCTGATCTTTTACCAATATATC 2040
Qy 2041 GGATAA 2046
Db 2041 GGGTAA 2046
```

```
RESULT 7
US-10-353-856-27
; Sequence 27, Application US/10353856
; Publication No. US20030215794A1
; GENERAL INFORMATION:
; APPLICANT: Kawasaka, Yoshihiro
; APPLICANT: Jasenosky, Luke D.
; APPLICANT: Neumann, Gabriele
; APPLICANT: Wisconsin Alumni Research Foundation
; TITLE OF INVENTION: Filovirus Vectors and No. US20030215794A1infectious Filovirus-Bae
; FILE REFERENCE: 800.032US1
; CURRENT APPLICATION NUMBER: US/10/353,856
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US 60/353,972
; PRIOR FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 19112
; TYPE: DNA
; ORGANISM: Marburg virus
US-10-353-856-27
```

```
Query Match 72.0%; Score 1473.2; DB 17; Length 19112;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 1688; Conservative 0; Mismatches 358; Indels 0; Gaps 0;
Qy 1 ATGAAGACCATATATTTCTGATTAGTCTCATTTTAAATCCAAAGTATATAAACTCTCCCT 60
Db 5940 ATGAAGACCATATGCTCTTTATCAGTCTTATCTTAATCCAGGATATAAACTCTCCCT 5999
```


; TITLE OF INVENTION: Development of a Preventive Vaccine for
; FILE OF INVENTION: Filovirus Infection in Primates

; CURRENT FILING DATE: 2004-03-26

; PRIOR APPLICATION NUMBER: PCT/US02/30251

; PRIOR FILING DATE: 2002-09-24

; PRIOR APPLICATION NUMBER: US 60/326476

; NUMBER OF SEQ ID NOS: 52

; SOFTWARE: RastSeq for Windows Version 4.0

; SEQ ID NO 31

; LENGTH: 7005

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic Construct pVR1012x/s Marburg GP (dTM)

US-10-491-121-31

Query Match 69.3%; Score 1418.8; DB 20; Length 7005;

Best Local Similarity 83.0%; Pred. No. 0;

Matches 1618; Conservative 0; Mismatches 332; Indels 0; Gaps 0;

Qy	1	ATGAAGACCATATATTTCTGATAGTCTCATTTTAATCCAAAGTATATAAACTCTCCCT	60
Db	2034	ATGAAGACCATATTTCTGATAGTCTCATTTTAATCCAAAGTATATAAACTCTCCCT	2093
Qy	61	GTCTTGAAGTGTCTAGTAAAGCCCACTCAAGATGTAGATTTCAGTGTCTCCGGAACC	120
Db	2094	ATTTTAGAGTAGCTAGTAAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAC	2153
Qy	121	CTCCAAAGACAGAGATGTTTCATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT	180
Db	2154	CTCCAGAGACAGAGAGCTCCATCTGATGGGATTTACATGAGTGGGCAAAAAGTTGCT	2213
Qy	181	GATTCCTCTTGGAGAGCATCTAAACGATGGGCTTTTCAGGACAGGTTTCCTCCCAAGAAC	240
Db	2214	GATTCCTCTTGGAGAGCATCTAAACGATGGGCTTTTCAGGACAGGTTTCCTCCCAAGAAC	2273
Qy	241	GTGAGTATACGAGGAGAGAGAGCCAAACATGTTACATATATAAGTGTACAGACCT	300
Db	2274	GTGAGTATACAGAGGAGAGAGAGCCAAACATGTTACATATATAAGTGTACAGACCT	2333
Qy	301	TCTGAAATTCCTGCTGCTGATCCTCCAGTAAATATCCGGAATACCTAAATGTAAA	360
Db	2334	TCTGAAATTCCTGCTGCTGATCCTCCAGTAAATATCCGGAATACCTAAATGTAAA	2393
Qy	361	ACTGTTTCATATATCAAGGTCAAAACCTCATGACAGAGGATTCCTCATTTGTGG	420
Db	2394	ACTATCCATCATATCAAGGTCAAAACCTCATGACAGAGGATTCCTCATTTATGG	2453
Qy	421	GGGCAATTTTCTGATGATCGCTGCTCTCAACAATGTACCGAGGCAAGTCTTC	480
Db	2454	GGAGCAATTTTCTGATGATCGCTGCTCTCAACAATGTACCGAGGCAAGTCTTC	2513
Qy	481	ACTGAAGGAAATATAGCAGCTATGATTTGTAATAGACAGATTACAGAAATGATTTTCT	540
Db	2514	ACTGAAGGAAATATAGCAGCTATGATTTGTAATAGACAGATTACAGAAATGATTTTCT	2573
Qy	541	AGGCAAGGACAGGTTATCGTCAATGAACCTTGACCTCCACCAATATATTTGCAAGC	600
Db	2574	CGGCAAGGACAGGTTATCGTCAATGAACCTTGACCTCCACCAATATATTTGCAAGC	2633
Qy	601	ACCAATGAACCGCAGAGAAATGATACGGGATGTTTGGATCTCCCAAGATACAACTCC	660
Db	2634	AGTAAACGGAACGCAACGAATGACACTGGAATGTTTGGGCGCTCTTCAAGAAATACAA	2693
Qy	661	ACAAACAAATCAAAATGCTCCCTCATCTCTTAAACCTCCATCCCTGCCACAGTAACCTCG	720
Db	2694	ACAAAGAACCAACATGCTCCCTGCTCCCAAAATACCTCCACCACTGCCACAGCCGCTCG	2753
Qy	721	AGCATTCATCTAACAATCTCAAAATTAATCTGCTAAATCTGGAACCTATGAACCAAGT	780

Db	2754	GAGATCAAACTCAACAAGCAACCCCAACTGATGCGACCAAACTCAATACCAAGGACCAAGC	2813
Qy	781	AGCGACGATGAGGACCTTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGCCCCACACA	840
Db	2814	AGTGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGAGAACGAGAACCCACACA	2873
Qy	841	ACTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	900
Db	2874	ACTTCTGATGCGGTCAACAAGCAAGGGCTTTTCATCAACAATGCCACCTCCCTCAGCA	2933
Qy	901	CATCCAAGCACTTCACACATGACCAAAAGTACGAATCTTCCCGCATGTCTGTAAT	960
Db	2934	CAACCAAGCAGCCACAGAGAGGAGAAACAACAACCAATTTCCCAAGATGCTGTGACT	2993
Qy	961	GAGCACAATGGAACCGACCCCAACCAACACACACACACACACACACACACACACACAC	1020
Db	2994	GAACCTAGACAAAATAAC	3053
Qy	1021	ACTCCCACTTAAACACTCTCAAGTACAACTCAAGTACAACTCAAGTACAACTCAAGTAC	1080
Db	3054	ATCTCTACTAAACACACTCCCAACACACAACTTCAGCACTCTCTCTGCAACCAATTA	3113
Qy	1081	ATCACCAATATGATATACACACGTTGAATAGCAGAAAGCGAACAACCAATGCTCAGTTG	1140
Db	3114	ACCACCAATGACAAACACACAGAGCAATCACTGAAATAGCAAAACACAGTGCCTCC	3173
Qy	1141	AACCAACTCTAGATCAACAGAGAAATCCCAACACAGACAGACACACACACACACACAC	1200
Db	3174	ATAACACCTGCTCTCCACCGGAAATCCCAACACAGACAGACACACACACACACACAC	3233
Qy	1201	AACATCATATGACGACATCATAGATATAAACAGCAAAACACACCAAAATTTCTCCGAT	1260
Db	3234	GGCCCCGCAACAGCGGACCAACACACAGCAATAGGCAATTTCCAGCTCTCCCTCCAC	3293
Qy	1261	TCTAGTCCGACAAACCGGCTCTTATATATCTTATAGTAAAGAAACGATCCATCTCTG	1320
Db	3294	CCCAGCTCGATGCAACCACTTTGATATTTTCAGAGAAAGCGAAGTATCTCTCTGAGG	3353
Qy	1321	GAAGGCACTGTTTCT	1380
Db	3354	GAAGGCACTGTTTCT	3413
Qy	1381	GTTCCAAATACAAAACAACTTTGATGAAATCTCTAGTCTGCTGCTCGCTCGAGAA	1440
Db	3414	GTTCCAAATACAAAACAACTTTGATGAAATCTCTAGTCTGCTGCTCGCTCGAGAA	3473
Qy	1441	GATCAACATGCTCCCTCCCAATATTTAGTAACTTTATCTTATTTTCTTAATAAATGAG	1500
Db	3474	GATCAACATGCTCCCTCCCAATATTTAGTAACTTTATCTTATTTTCTTAATAAATGAG	3533
Qy	1501	AACACTGCTACTCTCGAGAAATGAGAAATGATTTGATGAGAGTAAAGAAATTTGGAGC	1560
Db	3534	AACACTGCTACTCTCGAGAAATGAGAAATGATTTGATGAGAGTAAAGAAATTTGGAGC	3593
Qy	1561	GTTCAGGAGGATGACCTGGCGCAGGCTCAGTTGGATACCGTTTTTTTGGCCTCGAAT	1620
Db	3594	GTTCAGGAGGATGACCTGGCGCAGGCTCAGTTGGATACCGTTTTTTTGGCCTCGAAT	3653
Qy	1621	GAAGCACTTTACATGCTGTTTTTAAATTAATAATCAAAACAAATTTGGTCTCGAGGTT	1680
Db	3654	GAAGCACTTTACATGCTGTTTTTAAATTAATAATCAAAACAAATTTGGTCTCGAGGTT	3713
Qy	1681	CGTCTAGCAATCAAACTGCCAAATCTTTGGAATCTTTTATGAGAGTCAAACTCAGGAA	1740
Db	3714	CGTCTAGCAATCAAACTGCCAAATCTTTGGAATCTTTTATGAGAGTCAAACTCAGGAA	3773
Qy	1741	AGAACTTCTCTTAATCAATAGACATGCTAATGATCTTTTCTCAAGATGGGAGGA	1800
Db	3774	AGAACTTCTCTTAATCAATAGACATGCTAATGATCTTTTCTCAAGATGGGAGGA	3833
Qy	1801	ACATGCAAGGCTGGACCTGATTTGATCGGATAGAGACTTCTCCAAAAATATT	1860
Db	3834	ACATGCAAGGCTGGACCTGATTTGATCGGATAGAGACTTCTCTCCAAAAATATT	3893

Qy 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAACAAAGAGGGGACTGGTTGGGGTCTG 1920
Db 3894 TCAGAGCAAAATGACCAAAATTAATAAGGACGAACAAAGAGGGGACTGGTTGGGGTCTG 3953
Qy 1921 GGTGTTAAATGGTGACATCCGACTGGGGT 1950
Db 3954 GGTGTTAAATGGTGACATCCGACTGGGGT 3983

RESULT 9

US-10-491-121-32
; Sequence 32, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491,121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 8256
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct padapt Marburg GP (dtm)
US-10-491-121-32

Query Match 69.3%; Score 1418.8; DB 20; Length 8256;
Best Local Similarity 83.0%; Pred. No. 0;
Matches 1618; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

Qy 1 ATGAAGACCAATATATTTCTGATAGTCTCATATTTTAAATCCAAAGTATATAAACTCTCCCT 60
Db 1431 ATGAAGACCAATGTTCTCTTATCAGTCTTATCTTAAATTCAGGGGACAAAAAATCTCCCC 1490
Qy 61 GTTTTAGAATTTGCTAGTAACAGCCAACTCAAGATGTAGATTAGTGTGCTCCGGAC 120
Db 1491 ATTTTAGAGTAGCTAGTAATTAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAACT 1550
Qy 121 CTCAAAAAGACAGAGATGTTTCATCTGATGGATTACACTGAGTGGGCAAAAAGTTGCT 180
Db 1551 CTCAGAGACAGAGAGCTCCATCTGATGGATTACACTGAGTGGGCAAAAAGTTGCT 1610
Qy 181 GATTCCCTTTTGGAGAGCATATAACGATGGGCTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 1611 GATTCCCTTTTGGAGGATCCAAAGGATGGGCTTTCAGGACAGGTGTTACCTCCCAAGAA 1670
Qy 241 GTTGAGTATACGAGGAGAGAGCCAAACATGTTACATATATAGTGAACAGACCT 300
Db 1671 GTTGAGTATACAGAGGGGGAGAGCCAAACATGCTACAATATAGTGAACAGGATCCC 1730
Qy 301 TCTGAAAAATCTCTGCTGATCCTCCAGTAAATATCCCGGATTTACCTAAATGTAAA 360
Db 1731 TCTGAAAAATCTCTGCTGATCCTCTACCAACATCCCGTACTATCTTAATGCAAA 1790
Qy 361 ACTGTTTCATATATTCAGAGTCAAAACCTCATGACAGAGGAGTTGCCCTCCATTTGG 420
Db 1791 ACTATCCATCATATTCAGAGTCAAAACCTCATGACAGAGGAGTGCCTCTTCAITTAG 1850
Qy 421 GGGGCAATTTTCTGTATGATCGGCTTGCCTCTACAAATGATACCGAGGCAAGGTCTTC 480
Db 1851 GGAGCAATTTTCTGTATGATCGCATTTGCCTCCCAACAAATGATACCGAGGCAAGGTCTTC 1910

Qy 481 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAAGACAGTTTACAGATGATTTTTTCT 540
Db 1911 ACTGAAGGAAATATAGCAGCTATGATTTGTTAATAAGACAGTTTACAGATGATTTTTTCT 1970
Qy 541 AGCAAGGACAGAGTTATCTGTCATGAACTTGACCTCCCAATAATAATTTGACAGC 600
Db 1971 CGGCAAGGACAGAGTTATCTGTCATGAACTTGACCTCCCAATAATAATTTGACAGC 2030
Qy 601 AGCAATGAAACGACAGAGAAATGATA CGGGATGTTTGGCATCTCTCCAGAAATACAATCC 660
Db 2031 AGTAACGGAACGCAACGAAATGACACTGGATGTTTGGCGCTCTTCAAGAAATACAATCT 2090
Qy 661 ACAAACAATCAACATGCCCTCCATCTCTTAAACCTCCATCCCTGCGCCACAGTAACCTCG 720
Db 2091 ACAAAGAACCAACATGTCCTCCGTCCTCAAAATACCTCCACCACTGCCACAGCCCGTCC 2150
Qy 721 AGCATTCCTCTACAAATCTCAAAATTAATACTGTCTAAATCTGGAATCTATGAACCCAGT 780
Db 2151 GAGATCAAACTCAACAGACCCCAACTGATGTCACCAACTCAATACCGACCCCAAGC 2210
Qy 781 AGCGACGATGAGGACCTTATGATTTTCGGGCTCAGGATCTGGAGAACAGGGGCCCCACACA 840
Db 2211 AGTATGATGAGGACCTCGCAACATCCGGCTCAGGGTCCGGAGAACGAGAACCCACACA 2270
Qy 841 ACTCTTAATGATGACTGACGAAACAAATCGTCAACAATATTTGTCACATCTCTTCACTA 900
Db 2271 ACTTCTGATCGGTCAACCAAGGAGGCTTTTCATCAACAATGCCCACCACTCCCTCACA 2330
Qy 901 CATCCAGACCTCTACACATGAGCAAAACAGTAGTAAGTCTTCCCGACATGCTGTAACT 960
Db 2331 CAACAGCAGCCCAACAGAGGAGGAAACAAACAACCAATTTCCCAAGATGCTGTAACT 2390
Qy 961 GAGCACAATGGACCGGACCCCAACAAACAGCAACAGCAACGCTCTCTCAACAATACTAATA 1020
Db 2391 GAATGACAAATAATACAACTGACACACCGCTCCATGCCCCCTCATACACTTACCACA 2450
Qy 1021 ACTCCCACTTATTAACACTCTCAAGTACAACCTCAGTACTCTTCCCTCCCAACCCGCAAC 1080
Db 2451 ATCTCTACTAAACACACTCTCAACACCACTTACGACACTCTCTCTGACCACTTACAAAC 2510
Qy 1081 ATCAACCAATATGATACACACAGTGAAGTACGAGAAAGCAAAACCAATGCTCAGTTG 1140
Db 2511 ACCACCAATGACACACACAGAGCACAATCACTGAAAATGAGCAACACGAGTCCCTCCG 2570
Qy 1141 AACACAACTCTAGATCCAAACAGAAAATCCCAACACAGACAGCAACCAACAGCAACACC 1200
Db 2571 ATAACAACCTGCTCCAAACGGAATCCCAACACAGCAACAGCAACAGCAACAGCAACAAA 2630
Qy 1201 AACATCATATGACGACATCAGATATTAACAGCAACCAACCAACCAATTTCTTCCGAT 1260
Db 2631 GGCCCCGCCCAACAGGCAACCAACACAGCAAAATGAGCAATTTACAGCTCTCTCCGCCACC 2690
Qy 1261 TCTAGTCCGACCAACCCGCTCTCTATATATCTTTAGAAAGAAACGATCGATCTCTGGAG 1320
Db 2691 CCCAGCTCGATGCAACCACTTTGTATATTTTCAGAAAGAAACGAGATCTCTCTGGAG 2750
Qy 1321 GAAGCGACATGTTCCCTTTTCTGATGGGTTAATAATGCTCCAAATGATTTTGGACCA 1380
Db 2751 GAAGCGACATGTTCCCTTTTCTGATGGGTTAATAATGCTCCAAATGATTTTGGACCA 2810
Qy 1381 GTTCCAAATACAAAAACAATTTTGAATGAATCTCTAGTCTGTGCTCGGCTGAGGAA 1440
Db 2811 GTTCCAAATACAAAAACAATTTTGAATGAATCTCTAGTCTGTGCTCGGCTGAGGAA 2870
Qy 1441 GATCAACATGCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTAATAATAAGTAG 1500
Db 2871 GATCAACATGCTCCCAATATTTAGTTTAACTTTTATCTTATTTTCTTAATAATAAGTAG 2930
Qy 1501 AACACTGCTACTCTGAGGAAATGAGAAATGATTTGATGATGATGATGATGATGATGATGAT 1560
Db 2931 AACACTGCTACTCTGAGGAAATGAGAAATGATTTGATGATGATGATGATGATGATGATGATGAT 2990

```
QY 1561 GTTCAGGAGGATGACCTGGCCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAATT 1620
Db 2991 GTTCAGGAGGATGACCTGGCCGCGAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAATT 3050
QY 1621 GAAGGACTTTACACCTGCTGTTTTTAATTAATAAATCAAAACAATTTGGTCTGCAGGTGAGG 1680
Db 3051 GAAGGACTTTACACCTGCTGTTTTTAATTAATAAATCAAAACAATTTGGTCTGCAGGTGAGG 3110
QY 1681 CGTCTAGCCAAATCAAACTGCCAAATCCCTTGGAACTCTTATTGAGAGTCAACAACCTGAGGAA 1740
Db 3111 CGTCTAGCCAAATCAAACTGCCAAATCCCTTGGAACTCTTATTGAGAGTCAACAACCTGAGGAA 3170
QY 1741 AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTTACTCAACAAGATGGGAGGA 1800
Db 3171 AGAACATTTCTCTTAATCAATAGACATGCTATTGACTTTCTTACTCAACAAGATGGGAGGA 3230
QY 1801 ACATGCAAGTGTGGACCTGATTGTTGCATCGGGATAGAGACTTGTCCAAAAAATATT 1860
Db 3231 ACATGCAAGTGTGGACCTGATTGTTGCATCGGGATAGAGACTTGTCCAAAAAATATT 3290
QY 1861 TCAGAGCAAAATGACCAAAATTAATAAGGACGAACAAAAAGAGGGGACTGGTTGGGGTCTG 1920
Db 3291 TCAGAGCAAAATGACCAAAATTAATAAGGACGAACAAAAAGAGGGGACTGGTTGGGGTCTG 3350
QY 1921 GGTGTAATGGTGACATCCGACTGGGGT 1950
Db 3351 GGTGTAATGGTGACATCCGACTGGGGT 3380

RESULT 10
US-10-066-506A-5
; Sequence 5, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 5
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein 1
; OTHER INFORMATION: Marburg virus strain Raven Glycoprotein 2
US-10-066-506A-5

Query Match 64.5%; Score 1319.6; DB 15; Length 2046;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1592; Conservative 0; Mismatches 454; Indels 0; Gaps 0;

QY 1 ATGAGACCATATATTTCTGATTAGTCTCAATTTTAAATCCAAAGTATATAAACTCTCCCT 60
Db 1 ATGAAAGCCACATGTTTTCTTATCAGTCTTATCTTTAATTAATCAAGGACAAATAATCTCCCC 60
QY 61 GTTTTAAAGAAATGCTAGTAACAGCCAACTCAAGATGTAGATTTCAGTGTGCTCCGGAAACC 120
Db 61 ATTTTAGAGTAGCTAGTAATTAATCAACCCCAAAATGTGGATTGCGTATGCTCCGGAAT 120
QY 121 CTCAAAAAGACAGAAGATGTTCACTGATGGGAATTAACATGAGTGGGCAAAAAGTTGCT 180
Db 121 CTCAGAAGACAGAAGAGCTCCATCTGATGGGATTCACACTGAGTGGGCAAAAAGTTGCT 180
QY 181 GATTCCCTTTGGAGCATCTAAAGCATGGGTTTTCCAGACAGGTGTTCTCCCAAGAAC 240
Db 181 GATTCCCTTTGGAGCATCTCAAGCGTGGGTTTTTCAGGACAGGTGTACTCTCCCAAGAAC 240
```

```
QY 241 GTTGAGTATACGGAAGAGAAGAGCCAAACATGTTTCAATATAAGTGTAAACAGACCT 300
Db 241 GTTGAGTATACAGAGGGGAGGAAGCCAAACATGCTACAATAAAGTGTAAACGATCCC 300
QY 301 TCTGGAATTCCTTGTCTGCTCCAGTAAATATCCGGAATTAACCTTAATATGTAATA 360
Db 301 TCTGGAATTCCTTGTCTGCTCCAGTAAATATCCGGAATTAACCTTAATATGTAATA 360
QY 361 ACTGTTTCATCATATTTCAAGGTCMAAACCCCTCATGCACAGGGGATGCGCCTCCATTTTGTGG 420
Db 361 ACTATCCATCATATTTCAAGGTCMAAACCCCTCATGCACAGGGGATGCGCCTCCATTTTGTGG 420
QY 421 GGGGCATTTTCTTGTATGATCGGTTGCGCTCTACAAACAATGTAACGAGGCAAGGTCTTC 480
Db 421 GGAGCATTTTCTTGTATGATCGCATTTGCGCTCTCCACAAACAATGTAACGAGGCAAGGTCTTC 480
QY 481 ACTGAAGGAATATAGCAGCTATGATTTGTTAATAAGACAGTTTCACAGATGATATTTTCT 540
Db 481 ACTGAAGGGAACATAGCAGCTATGATTTGTTCAATAAGACAGTTTCACAGATGATATTTTCTCG 540
QY 541 AGGCAAGGACAAAGGTTATCGTCAATGAACCTTGACCTCCACCAATAAATATTGGACAAGC 600
Db 541 CGGCAAGGACAAAGGTTACCGTCAATGAATCTGACTTCTACTAATAAATATTGGACRAGT 600
QY 601 AGCAATGAAACGCGACAGAAATGATA CGGGATGTTTTTGGCATCTCTCCAGAAATACAACTCC 660
Db 601 AGTAAACGGAACGCAAAACGAATGACATCGGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 660
QY 661 ACAACCAATCAAACTGCCCTCCATCTCTTAAACCTCCATCCCTGCCACACAGTAACCTCG 720
Db 661 ACAAGAACCAAACTATGCTGCTCGTCCAAATACTCTCCACCTGCCCCACAGCCGCTCGG 720
QY 721 AGCAATTCATCTACAAATACTCATAAATACTCTAAATCTGGAACCTATGAACCCAAAGT 780
Db 721 GAGATCAAACTCAACAGCAACCCCAACTGTGTGECACCAAACTCAATACCAACGACCCCAAGC 780
QY 781 AGCGCATGAGGACCTTATGATTTTCGGGTCTAGGATCTGAGAACTGTGAGAACAGGGGCCCCACACA 840
Db 781 AGTGATGATGAGGACCTCGCAACATCTCGGCTCAGGGTCCGGAGAACAGGAAACCCCAACACA 840
QY 841 ACTCTTAATGTAGTCACTGAACAGAAACAACTCGTCAACAAATATTGTTCCACTCTTCACTA 900
Db 841 ACTTCTGATCGGTCAACCAAGCAAGGCTTTTCATCAACAAATGCCACCCACTCCCTCAACA 900
QY 901 CATCCAAGCACCTCAACAACATGAGCAAAAACAGTAGTAAGTCTCTCCGACATGCTGTAACT 960
Db 901 CAACCAAGCACGCCACAGCAAGGAGGAAACAAACAAACCAATTCCTCAAGATGCTGTGACT 960
QY 961 GAGCACAATGGAAACCGCAACCAACAAACAGCAAGCGTCTCTCAACAATACTAATAACA 1020
Db 961 GAACCTAGCAAAAAATAACAACTGCAACACCGTCCATGCCCTCTCATPAACACTACCACA 1020
QY 1021 ACTCCGACCTTATACACTCTCAAGTCAACCTCAGTACTCTCTCCCTCCCAACCGCAAC 1080
Db 1021 ATCTCTACTACAAACACTCCAAACACAACTTTCAGCACTCTCTGACCACTTACCAAAAC 1080
QY 1081 ATCAACCAATATGATACAAACGTTGAATAGCAGAAAGGCAACAAACCAATGCTCAGTTG 1140
Db 1081 ACCACCAATGACAAACACAGAGCACAACTCACTGAAATGAGCAAAACAGTGGCCCTCG 1140
QY 1141 AACCAAACTTCTAGATCCAAAGAAAATCCCAACCAAGGCAAGGACCAACAGCAACAAACC 1200
Db 1141 ATAACAACCTCTCCAAACGGGAAATCCCAACCAAGGACAGCAAGGACCAAGCAAGCAAAAA 1200
QY 1201 AACATCATGACGACATCAGATATTAACAGCAAAACCCCAACCAAAATTTCTTCCGGAT 1260
Db 1201 GGCCCCGCCCAACGGCACCACCAACAGCAAAATGAGCAATTTTCAGAGTCTCTCCGCCACC 1260
QY 1261 TCTAGTCCGCAAAACCCGCCCTCTATATATCTTTTAAAGAAAGAACGATCGATCTCTGGAGG 1320
Db 1261 CCCAGCTCGACTGCACACATCTTGTATATTTTCAGAGAAAGCGATCGATTTCTTGGAAA 1320
```

```
QY 1321 GAAGGACATGTTCCCTTTCTGATGGTTAATAATGCTCCAATTGATTTTGACCCA 1380
Db 1321 GAAGGTGATATATTTCCCGTTTTTTAGATGGGTTAATAATGCTCCAATTGATTTTGACCCA 1380
QY 1381 GTTCCAAATACAAAACAAATCTTTGATGAATCCTCTAGTCTCTGGTCCCTCGGCTGAGGAA 1440
Db 1381 ATCCAAACACAGAAACAAATCTTTGATGAATCCTCCAGCTTTAATACTTCAACTATGAG 1440
QY 1441 GATCAACATGCTCCCAATATATAGTTTAACTTTATCTTATTTTCTTAAATATAAATGAG 1500
Db 1441 GAACAAACACACTCCCGCCGAATATCAGTTTAACTTTCTCTTATTTCTGATAAAAAATGGA 1500
QY 1501 AACACTGCTACTCTGGAGAAAATCAGAAATGATGATGATGATGATGATGATGATGATGATG 1560
Db 1501 GATACTGCTACTCTGGGGAACACGAGAAATGATGATGATGATGATGATGATGATGATGATGAGT 1560
QY 1561 GTTCAGAGAGATGACTGCGCGCAGGGCTCAGTTGGATACCGTTTTTTGGCCCTGGAAAT 1620
Db 1561 GTGAGGAGGACGATTTGGCGCAGGGCTTAGCTGGATACCAATTTTGGCCCTGGAAATC 1620
QY 1621 GAAGGACTTTACCTGCTGTTTTTAAATTAATAATCAAAACAAATTTGGTCTGCAAGTTGAGG 1680
Db 1621 GAAGGACTTTACCTGCGGTTTTAATCAAAATCAGAAATTTAGTTTGTAGTTGAGG 1680
QY 1681 CGCTAGCAATCAAACTGCCAAATCCTTGGAACTCTTATTTGAGATCACAACCTGAGGAA 1740
Db 1681 CGCTTAGCTAATCAAACTGCTTAAATCCTTTGGAGCTCTTTTAAAGGTTCAACAACCGAGGAA 1740
QY 1741 AGAATCTTCTCTTAATCAATAGACATGCTAATGACTTTCTTACTCACAAGATGGGGAGGA 1800
Db 1741 AGGACATTTCTTAATCAATAGGATGCAATGACTTTTTGTCTTACGAGTGGGGGGA 1800
QY 1801 ACATGCAAGTGTAGGACCTGATTTGTCATCGGGATAGAGACTTTGTCCAAAAATATT 1860
Db 1801 ACATGCAAGTGTAGGACCTGATTTGTCATAGGAATAGAGATCTATCTAATAAATATC 1860
QY 1861 TCAGAGCAATTTGACCAAAATTAAGAGCAACAAAGAGGGGACTGGTTGGGGCTG 1920
Db 1861 TCAGAACAAATCGACAAATCAGAAAGGATGAACAAAGAGGAGAACTTGGCTGGGGCTA 1920
QY 1921 GGTGTAATGCTGGACATCGGCTGGGGTCTTCTTACTTACTTGGGATTTTCTACTA 1980
Db 1921 GGTGGCAATGCTGGACATCTGACTGGGGTGTCTCACCATTGTTGGGATCTCTGCTACTA 1980
QY 1981 TTATCCATAGCTGCTTGAATGCTCTATCTGTATTTGTCTGATCTTTTACTAATAATATC 2040
Db 1981 TTAATCTATAGCTGCTTGAATGCTCTGCTGCTATCTGCTGATCTTCACTAATAATATC 2040
QY 2041 GGATAA 2046
Db 2041 GGATGA 2046
```

RESULT 11

```
US-10-491-121-42
; Sequence 42, Application US/10491121
; Publication No. US20040259825A1
; GENERAL INFORMATION:
; APPLICANT: NABEL, GARY
; APPLICANT: YANG, ZHI-YONG
; APPLICANT: SULLIVAN, NANCY
; APPLICANT: SANCHEZ, ANTHONY
; TITLE OF INVENTION: Development of a Preventive Vaccine for
; TITLE OF INVENTION: Filovirus Infection in Primates
; FILE REFERENCE: NIH221.001NP
; CURRENT APPLICATION NUMBER: US/10/491.121
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/30251
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/326476
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SEQ ID NO 42
; LENGTH: 6902
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct pVR1012x/s Marburg (codon optimized)
US-10-491-121-42

Query Match      40.8%; Score 834; DB 20; Length 6902;
Best Local Similarity 64.3%; Pred. No. 1.3e-231;
Matches 1251; Conservative 0; Mismatches 695; Indels 0; Gaps 0;

QY 1 ATGAAGACCATATATTTTCTGATAGTCTCATTTTATCTCAAAAGTATATAAACTCTCCCT 60
Db 1923 ATGAAGACCATCTGCTGTTTCATCAGGCTGATCTCTGATCCAGGGCATCAAGACCTCGCC 1982
QY 61 GTTTTAAAGATTTGCTAGTAAACAGCAACCTCAAGATGTAGATTCAAGTGTCTCCGGAACC 120
Db 1983 ATCTTGAGATCGCCAGCAACACACGCCCCAGAACTGTGACAGCGTGTGAGCGGCACC 2042
QY 121 CTCCAAAGACAGAAAGATGTTCTATCTGATGGGATTTACACTGAGTGGGCAAAAAAGTTGCT 180
Db 2043 CTGCAGAAAGACCGAGGACGTGCACCTGATGGGCTTTCACCTGAGCGGCGCAGAGTGSCC 2102
QY 181 GATTTCCCTTTTGAAGATCTAAACGATGGGCTTTTCAAGACAGGTGTTCTTCCCAAGAAC 240
Db 2103 GACAGCCCTCTGGAGGGCAGCAAGAGGTGGGCTTTCAGGACCGGCGTGGCCGCCCAAGAAC 2162
QY 241 GTTCAGTATACGGAAGGAGAGAGCCAAACATGTTACATATATAGTGTAAACAGACCT 300
Db 2163 GTGAGGTACACCGAGGGGAGAGGAGGCAAGACCTTGCTCAACATCAGCGTGACCGACCCC 2222
QY 301 TCTGAAAAATCTTCTGCTGGATCTCCCAAGTAAATTCGCGATTAATACCTTAAATGTAAA 360
Db 2223 AGCGGCAAGGCTGCTGCTGGACCTCCCAACCAATCAGGGACTACCTTAAGTGCAG 2282
QY 361 ACTGTTTCAATATTTCAAGGTCAAAACCTCTATGCA CAGGGATTTGCTCCATTTGTGG 420
Db 2283 ACCATCCACACATCCAGGGCCAGAACCTCTCA CCGCCAGGGCATCGCCCTCAGCTGTGG 2342
QY 421 GGGCATTTTCTGTTATGATCGGTTGCTCTTACAAACATGTACCGAGGCAAGGTCTTC 480
Db 2343 GGGCCCTTCTTCTGTACGACAGGATCGCCAGCACCACTGTTACAGGGGGAGGGGTTC 2402
QY 481 ACTGAAGAAATATAGCAGCTATGATTTGTTAAATAAGACAGTTTCAAGAAATGATTTTTTCT 540
Db 2403 ACCGAGGCAACATCGCGCCATGATGTTTAA CAGACCGTGCACAGATGATCTTCAGC 2462
QY 541 AGCAAGGACAGGTTATCTCATGAACTTTGACCTTCCACCAATAAATATTGGACAGC 600
Db 2463 AGCGAGGCGCAGGGCTACAGGCACATGAA CCTGACACAGCAACCAAGTACTGGACCAGC 2522
QY 601 AGCAATGAACGAGAGAAATGATACGGGATGTTTGGCATCTCTCCAGAAATACACTCC 660
Db 2523 AACAAAGGCAACCCAGAGCAACAGACCGGCTGCTTGGCGCCCTGCGAGGAGTACACAGC 2582
QY 661 ACAAACAAATCAAACTCCCTTCCATCTCTTAAACCTCATCCCTCCCTGCCACAGTAACTCCG 720
Db 2583 ACCAAGAACAGACCTCGGCCCCAGCAAGATTC CCGAGCCCTTCCCAACCGCCAGGGCC 2642
QY 721 AGCATTCACCTTACAAATACTCAAAATTAATATCTGCTAAATCTGGAACCTATGAACCCCAAGT 780
Db 2643 GAGATCAAGCCCAACAGCACCCCAACGACGCAACCTTGAACACCAACCGACCCCAAC 2702
QY 781 AGCAGCATGAGGACCTTTATGATTTCCGGCTCAGGATCTGGAGAACAGGGGGCCCCACACA 840
Db 2703 AACGACGACGAGGACCTTGATCAC CAGCGGCGAGCGCGCGAGGAGGAGGAGCCCTACACC 2762
QY 841 ACTCTTAATGTAGTCACTGAACAGAAACAATCGTCAACAATATTGTTCCACTCTCTTCACTA 900
Db 2763 ACCAGGACCGCGTGACCAAGAGGGGCTGAGCAGGACCACTGCTCTTACCTTAGGCT 2822
QY 901 CATCCAAGCACCTCACAAACATGAGCAAAAAACAGTACGAATCTCTCCCGACATGCTGTAACT 960
```



```

Db 2823 CAGCCACGACCCCTCAGCAGGAGGCAACAACCGACACAGCGGACCGTACC 2882
Qy 961 GAGCAAAATGGAA CCGAGCCCAACAACAACAACAGCAACGCTCTCTCAACAATACTAATAACA 1020
Db 2883 GAGCCCAACAAGACCAACAACACCGCCAGCCAGCATGCTCTCTCAACAACACCAACGCC 2942
Qy 1021 ACTCCACCTATACACTCTCAAGTACAACTCTAGTACTCTCTCCCTCCCAACCGCAAC 1080
Db 2943 ATCAGCAACAACAACAGCAAGAAACAATCTTACGACCCCTGAGGGTGGAGCTGAGAAC 3002
Qy 1081 ATCACAATAATGATACACAACGCTGAACTAGCAGAAAGCGCAACAAACCAATGCTCAGTTG 1140
Db 3003 ACCACCACTAGCAACCCAGCAGCACCGCCACCGAGACGAGCAGACCGCCCTAGC 3062
Qy 1141 AACACAACCTTAGATCCAAAGAAATCCCAACCAAGCAAGACACCAACAGCAACCAAC 1200
Db 3063 AAGACCACTCTCCCAACCGCAACCTGACCAACCGCAAGAGCAGCAACAACACCAAG 3122
Qy 1201 AACATCATNTAGAGCATCAGATATACAGCAACCAACCCCAACCAATTTCTTCGGAT 1260
Db 3123 GGCCCCACCAACCGCCCTTAACATGACCAACGGCCACCTGACAGACCCCAAGCCCCACC 3182
Qy 1261 TCTAGTCCGACAACCGCCCTCTATATACCTTATAGAAAGAAACGATCGATCTCTGGAGG 1320
Db 3183 CCCAACCCCAACCCAGCAGCACTGTGTACTTTCAGGAAGAGAGGAGCATCTGTGGAGG 3242
Qy 1321 GAAGCGCATGTTCCCTTTCTTGGATGGGTTAATAAATGCTCCAAATGATTTTGACCCA 1380
Db 3243 GAGGGCGATATGTTCCCTTCTCGACGCGCTGATCAACGCCCTATCGACTGTCGACCCC 3302
Qy 1381 GTTCCAAATACAAAACAATCTTTGATGAATCTCTAGTCTGTGTCCTCGCGCTGAGGAA 1440
Db 3303 GTGCCCAACAACAAGACCATCTTCAGCAGAGCAGCAGCGCGCGCCAGCGCGAGGAG 3362
Qy 1441 GATCAACATGCTCCCAATATAGTTTAACTTTATCTTATTTTCCATAATAAATAGAG 1500
Db 3363 GACCAGCAGCGCCGCCCCAACATCAGCTGACCTGAGCTACTTCCCCAACATCAACGAG 3422
Qy 1501 AACACTGCTACTCTGAGAGAAATCAGAAATGATTTGATGAGAGTTAAGAAATTTGGAGC 1560
Db 3423 AACACCGCTACAGCGCGAGAACGAGAACGACTCGCAGCGGAGCTGAGGATCTGGAGC 3482
Qy 1561 GTTACGAGGATGACTCTGGCGCAGGGCTCAGTTGGATACCGTTTTTTTGGCCCTGGAAAT 1620
Db 3483 GTGCAGGAGGACGACTCGCGCGCGCTGAGCTGGATTCCTCTTCGCGCCCGGCATC 3542
Qy 1621 GAAGGACTTTACACTGCTGTTTAAATTAATTAATCAAAATTTTGGTCTGCAAGTTGAGG 1680
Db 3543 GAGGGCTGTACACCGCGCGCTGATCAAGAACCCAGAACCACTGGTGTGCAAGGCTGAGG 3602
Qy 1681 CGTCTAGCCAAATCAAACTGCAAAATCTTTGAACTCTTATTGAGAGTCAACAAGTGGAG 1740
Db 3603 AGGCTGGCCNACAGACCGCCNAGAGCTTGGAGCTGCTGAGGGTGACCAACCGAGGAG 3662
Qy 1741 AGAACTTCTCTTAATCAATAGACATGCTATGACTTTCTACTCAAGAATGGGGAGGA 1800
Db 3663 AGGACTTTCAGCTGATCAACAGGACACCGCATCGACTTCTGCTGACAGGTTGGGGCGGC 3722
Qy 1801 ACATCGAAAGTCTTGGACCTGATTTGTCATCGGGATAGAGACTTGTCCAAAATAAT 1860
Db 3723 ACCTCAAGAGTCTGGGCCCGACTGCTGATCGGCATCGAGGACCTGAGCAGGAAATC 3782
Qy 1861 TCAGAGCAAAATGACCAATTAATAAGGACCAAAAAGAGGGGACTGTTGGGGTCTG 1920
Db 3783 ACGGAGCAGATCGACAGATCAAGAAAGACGAGCAGAGAGGAGGACCGCGCTGGGGCCTG 3842
Qy 1921 GGTGTAATGTTGGACATCCGACTG 1946
Db 3843 GSCGCAAGTGTGGACCGGACTG 3868
```

RESULT 12

```

US-10-066-506A-3
; Sequence 3, Application US/10066506A
; Publication No. US20030108560A1
; GENERAL INFORMATION: Case C.
; APPLICANT: Grogan, Michael C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SAP
; CURRENT APPLICATION NUMBER: US/10/066,506A
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 3
; LENGTH: 1841
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Marburg virus strain Musoke Glycoprotein
; OTHER INFORMATION: Ebola virus Zaire Mayinga strain Glycoprotein 2
US-10-066-506A-3
```

```

Query Match 38.1%; Score 780; DB 15; Length 1841;
Best Local Similarity 74.8%; Pred. No. 3.4e-216; Mismatches 0; Gaps 0;
Matches 978; Conservative 0;

Qy 1 ATGAAGACCATAATATTTCTGATTAGTCTCATATTTTAATCCAAAGTATAAAAACTCTCCCT 60
Db 1 ATGAAGACCACATGTTTCTTATCAGTCTTATTTAATTCAGGGACAAAAAATCTCCCC 60
Qy 61 GTTTTGAATTTGCTAGTAACAGCAACCTCAAGATCTAGATTAGTGTGCTCCGGAAC 120
Db 61 ATTTTAGAGATAGCTAGTAATAATCAACCCCAAAATGTGGATTCGGTATGCTCCGGAAT 120
Qy 121 CTCCAAAGACAGAAAGTGTCTATCTGATGGGATTTACACTGAGTGGSCAAAAGTTTCT 180
Db 121 CTCGAAAGACAGAAAGCTCATCTGATGGGATTTACACTGAGTGGSCAAAAGTTTCT 180
Qy 181 GATTCCCTTTTGGAGCATCTAAACGATGGCTTTTCAGGACAGGTGTTCTCCCAAGAAC 240
Db 181 GATTCCCTTTTGGAGCATCCAAGGATGGCTTTTCAGGACAGGTGTTACCTCCCAAGAT 240
Qy 241 GTTCAGTATACGGAAGGAGAAAGCCAAACATGTTTCAATATATAAGTGTAAACAGCCCT 300
Db 241 GTTCAGTACACAGAGGGGAGAAAGCCAAACATGCTACATATAAGTGTAAACAGATCCC 300
Qy 301 TCTGAAAAATCTTGTCTGCTGGATCCCTCCAGTATATATCCGGATTAACCTTAAATGTAA 360
Db 301 TCTGAAAAATCTTGTCTGTTAGATCCCTCTACCAACATCCCGTACTATCCGAAATGCAA 360
Qy 361 ACTGTTCAATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATTCCTCCATTTGTGG 420
Db 361 ACTATCCATCATATTTCAAGGTCAAAACCTCATGCACAGGGGATTCCTCCATTTATGG 420
Qy 421 GGGGCATTTTCTTGTATGATCGGGTTCCTCTCAACAATGTATCCGAGGCAAGGTCTTC 480
Db 421 GGAGCATTTTCTGATATCGCATTCCTCCCAACAATGTATCCGAGGCAAGGTCTTC 480
Qy 481 ACTGAAGAAATATAGCAGCTATGATTTTAAATGAAGACAGTTTCAAGATGATTTTTTCT 540
Db 481 ACTGAAGGAAATATAGCAGCTATGATTTTAAATGAAGACAGTTTCAAGATGATTTTTTCT 540
Qy 541 AGGCAAGGACAGGTTATCGTCAATGAACCTTACCTCCCAACAATATAATTTGGACAAGC 600
Db 541 CGGCAAGGACAGGTTATCGTCAATGAACCTTACCTCCCAACAATATAATTTGGACAAGT 600
Qy 601 AGCAATGAACCGCAGAGAAATGATACGGGATGTTTTTGGCATCTCTCCAAAGATACAACTCC 660
Db 601 AGTAACGGAAACGCAACGAATGACATGATGTTTTCGGCGCTCTTCAAGAAATACAAATCT 660
Qy 661 ACAAAACAATCAAAATGCCCCTCCATCTCTTAAACCTCCATCCCTGCCCCACAGTAACCTCG 720
```

```
Db 661 ACAAGAACCAACATGCTCGTCCGTCCTCAAAATACCTCCACACCTGCCCCACAGCCCGTCCG 720
Qy 721 AGCAATCTCTCAAAATCTCAAAATCTCAAAATCTCAAAATCTCAAAATCTCAAAATCTCAAAAT 780
Db 721 GAGATCAAACTCACAGACCCCAACTGATGCCACCAAACTCAATACCAAGCCCAAGC 780
Qy 781 AGCAGCATGAGGACCTTATGATTTCCGCTCAGGATCTGAGATCTGAGACAGGGCCCCACACA 840
Db 781 AGTGATGATGAGGACCTTCGCAACATCCGGCTCAGGGTCCGGAGAACCCCAACACA 840
Qy 841 ACTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 841 ACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Qy 901 CATCAAGCACCTCAAACTGAGCAAAAGAGTACGAATCTTCCCGACATGCTGTAACT 960
Db 901 CAACCAAGCACCTCAAACTGAGCAAAAGAGTACGAATCTTCCCGACATGCTGTAACT 960
Qy 961 GAGCAATGGAACCGGACCCCAACACACACACACACACACACACACACACACACACACACACAC 1020
Db 961 GAATCTAGCAAAATAACACAACTGCAACCGTCCATGCCCTCTCAATAACACTACCA 1020
Qy 1021 ACTCCACCTAATACACTCTCAAGTACAACTCAGTACTCTTCCCTCCCAACCGCAAC 1080
Db 1021 ATCTCTACTAACAACCTCAAAACACAACTTTCAGCTCTCTCTGCAACCAATTAACAAC 1080
Qy 1081 ATCAACCAATATGATACAACTGTAAGTACGAGAAAGCGAACAACCAATGCTCAGTTG 1140
Db 1081 ACCCAATGACAAACACACAGACACAACTCACTGAATGAGCAACCCAGTGCCTCCG 1140
Qy 1141 AACCAACTCTAGATCCAAAGAAATCCCAACACAGACAGCAAGACACCAACAGACACAC 1200
Db 1141 ATAACAACTCTGCTCCAAAGAAATCCCAACACAGACAGCAAGACACCAACAGACACAA 1200
Qy 1201 AACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1201 GGCCCGGCAACAGCGCAACCAACAGCAAAATGAGCAATTCACAGTCTCTCCCGAC 1260
Qy 1261 TCTAGTCCGACACCCGCTCTCTATATATCTTGAAGAAAGAACGATCG 1308
Db 1261 CCCAGCTCGACTGCACACAACTCTGTATATTTTCAGAAAGAGCGATCG 1308

RESULT 13
US-10-066-506a-1
; Sequence 1, Application US/10066506a
; Publication No. US20030108560A1
; GENERAL INFORMATION:
; APPLICANT: Grogan, Case C.
; APPLICANT: Hevey, Michael C.
; APPLICANT: Schmaljohn, Alan, L.
; TITLE OF INVENTION: Chimeric Filovirus Glycoprotein
; FILE REFERENCE: 003/243/SNP
; CURRENT APPLICATION NUMBER: US/10/066,506a
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/267,522
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 1
; LENGTH: 2252
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chimeric molecule between Ebola virus Zaire Mayinga strain Glycoprotein and Marburg virus strain Musoke Glycoprotein 2
US-10-066-506a-1
Query Match 36.5%; Score 746.6; DB 15; Length 2252;
Best Local Similarity 99.5%; Pred. No. 2.1e-206;
Matches 749; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 1294 AGAAAGAAACGATCGATCTCTCGAGGGAAGCGACATGTTCCCTTTCTGGATGGGTGA 1353
Db 1492 AGAACTCGAGATCGATCTCTCGAGGGAAGCGACATGTTCCCTTTCTGGATGGGTGA 1551
Qy 1354 ATAAATCTCCAAATTTGATTTTGAACCCAGTTCCAAATACAAAAACAATCTTTGATGAATCC 1413
Db 1552 ATAAATCTCCAAATTTGATTTTGAACCCAGTTCCAAATACAAAAACAATCTTTGATGAATCC 1611
Qy 1414 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1473
Db 1612 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1671
Qy 1474 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1533
Db 1672 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1731
Qy 1534 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1593
Db 1732 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1791
Qy 1594 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1653
Db 1792 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1851
Qy 1654 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1713
Db 1852 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1911
Qy 1714 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1773
Db 1912 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1971
Qy 1774 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1833
Db 1972 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 2031
Qy 1834 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1893
Db 2032 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 2091
Qy 1894 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 1953
Db 2092 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 2151
Qy 1954 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 2013
Db 2152 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 2211
Qy 2014 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 2046
Db 2212 TTAGTTCTGGTCTCGGCTGAGGAAGATCAACATGCTCTCCCAATATATTAGTTTAACT 2244

RESULT 14
US-10-353-856-9
; Sequence 9, Application US/10353856
; Publication No. US20030215794A1
; GENERAL INFORMATION:
; APPLICANT: Kawaoka, Yoshihiro
; APPLICANT: Jasenosky, Luke D.
; APPLICANT: Neumann, Gabriele
; APPLICANT: Wisconsin Alumni Research Foundation
; TITLE OF INVENTION: Filovirus Vectors and No. US20030215794A1
; FILE REFERENCE: 800.032US1
; CURRENT APPLICATION NUMBER: US/10/353,856
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: US 60/353,972
; PRIOR FILING DATE: 2002-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 18890
; TYPE: DNA
```

